



```

170 TyrGluLysAlaGlyValAlaProPheAlaGluGluAspAspProAlaGlu 186
531 ATACGGCTGCGTTTGTGTACACTCAAGCAGGCTTCAACCAAGACAGAGA 580
186 uTrpLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLysA 203
581 ACCATGGCAGACAGCTGCTGCGGCGATGAAACGCTTTTCCGAAGACACA 630
203 sPlsYStrPThrLysIleAlaGluSerValLysGlyValThrGluGluThr 219
631 ACAACAGGTGTCACCGCTTACAGCTCCGAGAGAGAGAGAGAGAGAGAGAG 680
220 ThrThrGlyValLeuArgLeuArgLeuArgLeuAlaAlaGlyAspLeuAl 236
681 CTTCACAGCCATCAACGTCACAGAGCTGTTCACAAAGTCCAAAGTCCGATA 730
236 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 253
731 ACATCTACGGCTGCGGCGCATCTTCATGATGATGATCAACGCTGCTTCC 780
253 snLysTYrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 269
781 GATGTCATGATCGGCGGCAAGACAGCTCTGTCATGAGGTTCAGCGCATGT 830
270 AspAlaLeuIleGlyGlyLysLysValLeuIleCysGlyTYrGlyAspVa 286
831 CGGCAAGGGCTGCGCTCAATCCCTCGTGGCAAGGCGCTCGGCTATCA 880
286 lGlyLysGlyCysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 303
881 TCACAGAACTGACCCATCTGCGCTCTCCAGAGCTCCATGAGAGCTTAC 930
303 aThrThrIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 319
931 CAGGTCCGCGCATCGAGGAGAGTGGCAAGATGTCATGATATCTTCCTTAC 980
320 AspValValThrValGluGluAlaIleGlyAspAlaAspLeuValAlaTh 336
981 ATGCACAGAAACTGCGATATCATCTGTGTACATGATGATGCGCCAGATGA 1030
336 rAlaThrGlyLysLysAspIleIleMetLeuGlnHisIleLysAlaMetL 353
1031 AGGATTAAGGCTATGTGCGTAAACATCGCCACTTGATTAACCAATTGAT 1080
353 ySAspRHisAlaIleLeuGlnLysIleGlyHisPheAspAsnGluIleAsp 369
1081 ACAGATGGCTCATGAATATACCCAGCATCAAGCATCCCAATCAAGCC 1130
370 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysP 385
1131 AGAATACGACATGTGGGAATTCAGAT...GGCCAGGCTATCCTCTTC 1177
385 oGlnValAspLeuThrPheGlyAspThrGlyArgSerIleIleValL 402
1178 TTGGTGAAGGCGGCTTCTTAACCTTGGCTGGCTACAGAGTCAACCATCT 1227
402 euSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 418
1228 TTGGTTATGTCAATGTCTACAAACAGACACTGCTCAGCTCGACCT 1277
419 PheAlaMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 435
1278 CTACGAAAGAGAGGAAATCTCGAGAGAGAGGTTTACACACTTCCGAAAC 1327
435 uTrpThrLysAsnAspGluTYrAspAsnGluValTYrArgLeuProLysH 452
1328 ATCTCGATGAAGAGCTGCGTCCGCTCCACCTCGGATCTTCGATGTCAC 1377
452 lSLeuAspGluLysValAlaIleArgIleHisValGluAlaLeuGlyLysH 468
1378 CTTACAAAGCTTACACAGAGCAGGCTGACATCAACGCTTCCAGTTTGA 1427

```

```

469 LeuThrLysLeuThrLysGluGlnAlaGluTYrLeuGlyValAspValAl 485
1428 GGTCCTTACAGTGTGATGCTTACGCTTAT 1458
485 uGlyProTYrLysProAspHisTYrArgTYr 495

```

seq\_name: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:US-09-708-427-22304

seq\_documentation\_block:

; Sequence 22304, Application US/09708427

; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEP

; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427

; NUMBER OF SEQ ID NOS: 85364

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 22304

; LENGTH: 485

; TYPE: PR

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 1..485

; OTHER INFORMATION: xaa is any amino acid

; NAME/KEY: misc\_feature

; LOCATION: 1..485

; OTHER INFORMATION: Ceres Seq. ID 1840722

; US-09-708-427-22304

alignment\_scores:

Quality: 1537.00 Length: 487

Ratio: 3.901 Gaps: 4

Percent Similarity: 80.903 Percent Identity: 62.012

alignment\_block:

US-09-759-990-1 x US-09-708-427-22304 ..

Align seg 1/1 to: US-09-708-427-22304 from: 1 to: 485

```

19 ACTGTCGTCATTCGATGACAGAAATTCGACATCAACCTCCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluTYrLysValLysAspMetSerGlnAlaAspH 24
69 CGGCGGTAAAGAACTTACCTTCTGTGAGAAAGAAATTCGAGCTTTATGG 118
24 eGlyArgLeuGluLeuGluLeuAlaGluValAlaGluMetProGlyLeuMet 41
119 TTCTTCGTGAGAGGTTATTCGCTTCTTAAGCCATTGAAGGTCAGAGAT 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 lacyArgThrLurPheGlyProSerGlnProPheLysGlyAlaIle 57
169 TCTGTTCCCTCCACATGACAGTCCAGACAGCCGCTTCATCGAGACAT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrIle 74
219 CACACCTCTGTGCTGATGATGATGAGGCTTCTTCGACATCTTCTTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTYrPheLysSerCysAsnIlePheSer 91
269 CACAAGATACAGCCGCTGCTGATGATGATGTCGCGCCCAACGAGACACA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 hrgLysAspHisAlaAlaAlaIleAlaIleArgAspSerAlaAla..... 105
319 GAGAGCCAGCCGCTATCCCACTCTTCGCTGAGAGGCGCAACACATCCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTrpLysGlyLurThrLeuGlu 115
369 AGAATACGGGAGAACATACCGGCTCTCACATGCGCCAGATGTCAG 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

115 nglutrrprpcythrghluaralaleuasprpglyproglylylg 132
419 GCCACACAGAGTTCGATGATGGTGTGATGCTACATCCATCC 468
132 lpproaspheulevalaspasglylyaspalathleuileuileis 148
469 AAGGCTTCGAA.....TTCGAAACAGCCGCTGCTGCTCAGA 506
149 gluglyvalylaglugluilephegllulysrghlylnvalproas 165
507 GCCACACAGAGTTCGATGATGGTGTGATGCTACATCCATCA 556
165 pprothserthrpspnprogluphegllilevalleuilerleilel 182
557 AGCAGGTCTCAACCAAGACACAGACACAGCAGTTCGCTCCGCG 606
182 ysgluglyleuglnvalaspprolylslyrhlslslysmetlysluar 198
607 ATGACGCGTGTTCGAGAGACACACAGTTCACCGCTTCACCA 656
199 leuvalglyvalsergluutrrthrthrglyvalylsarglyutrgl 215
657 GCTCGAAGAGAGGCAACTCCTCTTCCAGCATCAACGTCACAGAG 706
215 nmetglnghasnghlythrleuupheproalaleasnvalasnaps 232
707 CTGTACAAGTTCAGTTCATCAATCTACGCGCTCCGCTCCCT 756
232 etvalthrlysserlyspheaspnsneutryglycysarghlsserleu 248
757 ATCGATGTATCAACCGTCTCCGATGTATGATGGCGGCAAGACAC 806
249 proaspglyleumetaralathraspvalmetlealaglylyvalal 265
807 TCTCGCATAGGTTTACGCGCATGTCGCAAGGCGCTCATCCCTCC 856
265 aavalliecygsltyrghlyaspvalglylygslasalaalalemetl 282
857 GTGGCCAAAGGCGCTCGCTTATCATCAACAGCAACTGACCCAA 906
282 ysthrthraglyalargvallevalthrghluileasproillecysala 298
907 CTCGAGGCTGCGCATGAGAGGCTTACAGGTCGCGCATGAGAGAA 956
299 leuglnalaleumetgluglyleuglnvalleuthrleuqluaspa 315
957 CAAGGATGCGATATCTTCCTTACATGACAGAAACTGCGATATCAT 1006
315 lserglualasphevalthrleuthrthrglyasnlysa 332
1007 CTGTTGATGATGAGCCAGATGAGATGAGATGAGTATGTCGGTAA 1056
332 etvalasphevalarglymetlysaasnalevalcysasalle 348
1057 GCGCAGCTTCGATGAGAAATGATACAGATGCGCTCATGAAATACC 1106
349 glyhispheaspasnghluileaspmetleuglyleuglnuthrly 365
1107 CATCAACAGCACTCCCAATGAGCCAGATGAGATGAGATGAGAT 1156
365 yvalylsarglthrleuthrleuthrproglthrspargtrpvalph 382
1157 ATGGCCAGCT...ATCCCTCTTCTGCTGAGGCGCGCTCTTAACT 1203
382 luthrlysalaglylellevalleualagluglyargleumetasnleu 398
1204 GCGCGCTTCAGGTCACCACTCTTCTGATGATGATGATGATGAT 1253
399 glycysalathrghlyhispheasphevalmetsercysserp 415
1254 CAGACAGCTCGCTGAGCTGAGCTGAGCAAGAAAGAGA.....GGA 1297
415 nglvalleleaglnleuglnleuthrpsnghlysalaserghlyst 432

```

```

1298 TCGAAGAGAGGTTTACACACTTCGAGCATCTCGATGAAAGACTCCT 1347
432 yrglulyslysaltyrvalleuprolyshisleuaspglulysvala 448
1348 CGCCTCCACCTCGGATCTCTCGATGTCACACTTCGAAAGCTTCAC 1397
449 leuileuileuileuileuileuileuileuileuileuileuileu 465
1398 GCAGGCTGACTATCATCAACGTCGATGAGAGGTCCTTACAGTGT 1447
465 pelnseraptyrvalserleuprolleugllyprotyrlysp 482
1448 CTTACCGTTAT 1458
482 lstryatgyr 485

seq_name: /cgn2_6/ptodata/2/pae/US60_COMB.pep:US-60-361-294-16

seq_documentation_block:
; Sequence 16, Application US/60361294
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua
; APPLICANT: McElver, John Alan
; APPLICANT: Zhou, Qing
; APPLICANT: Aux, George W.
; APPLICANT: Tossberg, John
; APPLICANT: Ashby, Carl
; APPLICANT: Dunn, Jill
; APPLICANT: Cates, Eddie
; APPLICANT: Law, Marcus Dixon
; APPLICANT: Budzieszewski, Greg
; APPLICANT: Nyfeler, Beat
; APPLICANT: Wu, Jia-Qian
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential For Pla
; FILE REFERENCE: 60169Pl
; CURRENT APPLICATION NUMBER: US/60/361,294
; CURRENT FILING DATE: 2002-02-28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 485
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
US-60-361-294-16

alignment_scores:
Quality: 1537.00 Length: 487
Ratio: 3.901 Gaps: 4
Percent Similarity: 80.903 Percent Identity: 62.012

alignment_block:
US-09-759-990-1 x US-60-361-294-16 ..
Align seg 1/1 to: US-60-361-294-16 from: 1 to: 485

19 ACTGCTGCTCATTCGAGTACAGAAATGGCCGATCAACCTCATGTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 Thrserserglyargglutrylsvalylsaspmetserscinalaasph 24
69 CGGCGGTGAGAACTTACCTTGCCTGAGAAAGAAATCCGAGCTTATGG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 egllyargleuglnleuglnleualagluvalglumetp 118
119 TCTCTGAGAGGTTATTCGCTTCTTACGATGAGAGGTCGATGATC 168
::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 lacysarqfthcluhbeglyproserclnproheylsglyalargyle 57
169 TCTGCTCCCTCGATGACAGTCCGAGACAGCGCTCTCATGAGACACT 218
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
219 CACAGCTTTGGTGTGATGATGATGGGCTTCTGCAACATCTTCTCTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGlyValAlaGlyTrpCysSerCysAsnIlePheSer 91
269 CACAGATACAGCGGCTGCTGATGCTGTGTGCGGCCAACAGGACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 hGlnIleSphIleAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319 GAGAAAGCCAGCGGTATCCAGTCTTCCGAGGAGGCGCAACACCTCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaIleTrpLysGlyIleuThrLeuGlu 115
369 AGAATACTGGGAGAACATACCCTGCTCACAATGGCCAGATGCTCAAG 418
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 nGluTrpTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGly 132
419 GCCCAGACAGGTTGTCGATGATGATGATGATGATGATGATGATGAT 468
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 LyrProAspLeuIleValAspAspGlyAlaAspAlaThrLeuIleHis 148
469 AAGGCTTCGAA.....TTCGAAACAGCGGCTGCTTCGAGA 506
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 GluGlyValIleGluGluIlePheGluLysThrGlyGlnValProAs 165
507 GCCAACAGAGCTGACCAACCTGGAATACCGCTCGTTTGTCTACACTCA 556
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIle 182
557 AGCAGGTTCTCAACACAGACAGACACACACACAGTTGCTGCCGCGC 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 YsgLuglyLeuGlnValAspProLysLysThrIleLysMetLysGluArg 198
607 ATGAACGGTGTTCGGAAGAGACAAACACAGGTGTCACCGCTTACCA 656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTrp 215
657 GCTGAGAGAGGAGGCAACTCTCTCCACGACATCAACGTACAGACG 706
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 nMeGlnIleAsnGlyThrLeuLeuPheProAlaIleAsnValAspSps 232
707 CTGTTCACAACTCCAGTTCGATACATCTACGCGTGGCGGCTGCTT 756
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 eValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 248
757 ATCATGATGATCAACCGTCTTCGATGATGATGATGATGATGATGAT 806
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyValAla 265
807 TCTGTCATGGGTTACGGGATGTCGCAAGGCGGCTGCAATCCCTCC 856
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 aValIleCysGlyTrpGlyAspValGlyLysGlyAlaAlaAlaMet 282
857 GTGGCCAGAGCGCTGCGCTTATCATCACAGACTGACCCCAATTCGCT 906
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 YsThrAlaGlyAlaArgValIleValIleThrGluIleAspProIleCysAla 298
907 CTCCAGGCTGCCATGGAAGGTACAGGTCCGCCCATCGAGGAAGTCT 956
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 LeuGlnIleLeuMetGluGlyLeuGlnValLeuThrLeuGluAspVal 315
957 CAAGATGTCGATATCTGTTACATGCAAGAAATGCGCATATCATCT 1006
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 LserGluAlaAspIlePheValIleThrThrGlyAsnLysAspIleIle 332
332 eValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
1057 GGCCACTTCGATTAAGAAATGATACAGATGCTCATGAATACCCAG 1106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGlyThrTrpProG 365

```

```

1107 CATCAAGCACATCCCAATCCAGCCAGAAATACAGATGTGGATTTCCAG 1156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 YValLysArgIleThrIleLysProGlnThrAspTrpValPheProG 382
1157 ATGGCCAGCT...ATCCTCCTTCTGCTGAGAGGCGGCTTTCATCTT 1203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 LuThrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
1204 GCGTGGCTTACAGCTCACCCATCTTTCGTTATGTCATGTCATCAAA 1253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThr 415
1254 CCAGACCTGCTCAGCTCGACCTTACGAAAGGA.....GGAATTC 1297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 nGluValIleAlaGlnLeuGluLeuIleTrpAsnGluLysAlaSerGly 432
1298 TCAGAGAAGAGTTTACACACTTCCGAGATCTCGATGATGAGAAATGCT 1347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 YrGluLysLysValIleValLeuProLysHisLeuAspGluLysValAla 448
1348 CGGCTCCAGCTCGATCTCTGATGTCCACCTTACAAAGCTTACACAGA 1397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLys 465
1398 GCAGGCTGATCATCAACGTTCCAGTTGAGGTCCTTACAAAGTGTGATG 1447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465 pGlnSerAspTrpValSerIleProIleGluGlyProTrpLysProPro 482
1448 CTTTACCGTTAT 1458
|||||:|||||:
482 LsTrpArgTrp 485

```

seq\_name: /cgn2\_6/ptodata/2/pae/US097\_COMB.pep:US-09-708-427-77520

```

seq_documentation_block:
? Sequence 77520, Application US/09708427
? GENERAL INFORMATION:
? APPLICANT: N. ALEXANDROV et al.
? TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEP
? FILE REFERENCE: 2750-1243P
? CURRENT APPLICATION NUMBER: US/09/708,427
? CURRENT FILING DATE: 2000-11-09
? NUMBER OF SEQ ID NOS: 85364
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 77520
? LENGTH: 485
? TYPE: PRT
? ORGANISM: Zea mays subsp. mays
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 1..485
? OTHER INFORMATION: Xaa is any amino acid
? NAME/KEY: misc_feature
? LOCATION: 1..485
? OTHER INFORMATION: Ceres Seq. ID 1960872
? US-09-708-427-77520

```

alignment\_scores:

Quality:	1523.00	Length:	487
Ratio:	3.875	Gaps:	4
Percent Similarity:	80.698	Percent Identity:	61.807

alignement\_block:

US-09-755-990-1 x US-09-708-427-77520 ..

Align seg 1/1 to: US-09-708-427-77520 from: 1 to: 485

```

19 ACTGATGCTCATTCGAGTACAGAAATGCGCAACATCACTCATTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerGlyArgGluTrpLysValLysAspLeuSerClnAlaAsp 24

```



```

69 CGCGCGTTAAGCACTTACCCCTTCTGAGAAAGAAATGCCAGCTTTATGG 118
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
24 eglYArleuGlulIeGlIuLeuValGluMeCProGlyLeuMeTa 41
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 TTTCTCGAGGGTTATCCGCTTCTAAGCACTTGAAGGGTGTGACAGATTC 168
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 IAcYArGAlaGluPheGlyProSerIysProPheAlaGlyAlaArgIle 57
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 TCTGTCTCCCTCCACATGACAGTCACAGACGCGCTCCATCGAGACACT 218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrIle 74
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 CACAGCTCTGTGCTGATGATGATGAGTGGCTCTCCGACATCTTTCTTA 268
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerT 91
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 CACAGATACACAGCGCTGCTGATGCTGTGCGGCCACAGCGACACCA 318
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 hrGlnAsPHisAlaAlaAlaAlaIleAlaArgSPSerAlaAla..... 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 GAGAAGCCAGCGGTATCCAGTCTTCCGCTGAGAGGGCGAAGACACTCC 368
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 .....ValPheAlaTrpIlySglYgluThrLeuGl 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 AGATACTGGGAGAACACATACCAGCGCTGCACATGCGCCAGATGGTCAAG 418
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 uGluTrpTrpCysThrGluArgCysLeuAspTrpGlyGluAlaGlyG 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 GCCCAGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGCC 468
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 IyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
469 AAGGGCTTC.....GAATTGAAACAGACCGGCTGCTGTCCAGA 506
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 GluGlyValIlySAlaGluGluGluGluGluGluGluGluGluGluGlu 165
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
507 GCCACAGACAGCTGACACCTGACATACCGCTGCTTCTGTCATCACTCA 556
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 pProGluSerThrAspAsnAlaGluPheIlySleValLeuThrIleIleA 182
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
557 AGCAGGCTCTCAACCAAGACCAAGACCAAGACCAAGTGTGCGCCGCG 606
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 rGaSPGlyLeuIlySAlaAspProIlySlyTrArgIlySMeIlySAlaArg 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
607 ATGAACGGTGTTCGAGAGAGACAAACAGAGTGTCCACGCGCTTACCA 656
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 LeuValGlyValSerGluGluThrThrThrGlyValIlySArgLeuTrpGl 215
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
657 GCTCGAGAGAGGCGCAACTCCTCTCCAGCCATCAACGTCACGACG 706
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 meIcIInGluThrGlyAlaLeuPheProAlaIleAsnValAsnAspS 232
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 CTGTTCACAAGTCCAAAGTTCGATACATCTACGCGTCCGCCACTCCCTT 756
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 eValThrIlySerIysPheAspAsnLeuTrpGlyCysArgHisSerLeu 248
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
757 ATCGATGATACACCGCTCTCCAGTGCATGATCGCGCGCAGACAGAC 806
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyIlySValAl 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 TCTGCTCATGGGTTCAGCGGATGTCGGCAAGGCTGCGCTCAATCCCTCC 856
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 aValValAlCysGlyTrpGlyAspValGlyIlySAlaAlaAlaLeuI 282
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
857 GTGGCCAGAGGCGCTGCGCTTATCATCAAGACGACGACGACGCGCT 906
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 ySAlaAlaGlyAlaArgValIleValThrIleIlySProIleCysAla 298
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
907 CTCAGAGCTGCATGAGAGGCTACAGGTCGCGCCATCGAGAGAGTGT 956
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuProLeuGluAspValIva 315

```

```

957 CAAGATGTCGATATGTTCTTGCATGACACAGGAACCTGCCATTCATCT 1006
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 IserGlnAlaAspIlePheValThrThrThrGlyAsnIlySAspIleIleM 332
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1007 CTGTTGACATGATGGCCAGATGAAGGATGAAGGCTATGTGGTAAACATC 1056
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 eValAspHisMetArgIlySMeIlySAsnAlaIleValAlCysAsnIle 348
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1057 GGCACCTTCGATACGAAATTTGATACAGATGGCCCTCATGAATACCAAG 1106
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
349 GlnHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTrpProGl 365
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1107 CATCAAGCACATCCCAATCAAGCCAGAAATACAGATGTGGGATTTCCAG 1156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 YAllySArgIleThrIleIlySProGlnThrAspArgTrpAlaPheProG 382
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1157 ATGGGCAC...GCTATCCTCTTCTGAGGGCGCGCTTCAACCTT 1203
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 IuThrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1204 GCGTCGCTTACAGGTACACCATGTTCTGTTATGTCATGTCATTCACAAA 1253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1254 CCAGACACTGCTCAGCTCGACCTTACGAAAGA.....GGAATC 1297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415 nGlnValIleAlaGlnLeuGluLeuTrpIlySgluIlySerSerGlyLysT 432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1298 TCGAGAGAGAGTTTACACACTCCGAAAGCATCTCGATGAGAGAACTGCT 1347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
432 yGluIlySlySValTrpValLeuProIlyShIleAspGluIlySValAla 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1348 GCGCTCCACCTGGATCTCTGATGTCCACCTTACAAAGCTTACACAGA 1397
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
449 AlaLeuHisLeuGlyIlySLeuGlyAlaIlySLeuThrIlySLeuThrIlyS 465
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1398 GCAGGCTGACTATCATCAAGCTTCCAGTTGAGAGGTCCTTACAGTGTATG 1447
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 rGlnAlaAspTrpIlySerValProIleGluGlyProTrpIlySProAlaH 482
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1448 CTTACCGTTAT 1458
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
482 IStYTrArgTrpTyr 485
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq\_name: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:us-09-708-427-77519

seq\_documentation\_block:

```

; Sequence 77519, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPE
; FILE REFERENCE: 2750-1243p
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77519
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..510
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..510
; OTHER INFORMATION: Ceres Seq. ID 1960871
; US-09-708-427-77519

```

alignment\_scores:

Quality: 1523.00

Length: 487

Ratio: 3.875 Gaps: 4  
Percent Similarity: 80.698 Percent Identity: 61.807

Alignment\_block:  
US-09-759-990-1 x US-09-708-427-77519 ..

Align seg 1/1 to: US-09-708-427-77519 from: 1 to: 510

```

19 ACTGTCCTCCATTCGAGTACAGATTCGCGACATCAACCTTCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 Thrserserlglyargglutrylserlyvalylsaspleserlnalasp 49
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 CGGCCCGTAAGAACTACCTGCTGAGAAAGAAATGCCAGCTTTATAG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
49 eglvalgluenglilegluleuvalgluvalgluvalgluvalglu 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
119 TTTCTGTCGAGCGTTATTCCTCTTACAGCATTAAGAGGCTGAGATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 lacsarlgalaglupheglyproserlysprophelaglyalarglie 82
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
169 TCTGTCTCCCTCCACATGACATGACAGCCGCTCTCATGAGACACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
83 SerglyserleuHismethrileglthrAlaValleuileglutHrie 99
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
219 CACAGCTTGTGTCATGTCAGATGGGCTTCTGCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
99 utHrAlaLeuGlYAlaGlUValArgTrpCysSerCysAsnIlePheSer 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
269 CACAGATACAGCGCTGCTATCGTTGCGGCCCAAGGCAACAGCAACA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
116 hrGlmsphAlaAlaAlaAlaAlaAlaAlaArgSPSerAlaAla.... 130
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
319 GAGAAAGCAGCCGGTATCCAGTCTGCTGAGAGGCGGAACAACCTCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
131 .....ValPheAlaTrpLysGlyLutHrLeuGl 140
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
369 AGAATACTGGGAGAACACATACCGCGCTCTCATAGGCCAGATGGTCA 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
140 uGlutYTrpTrpCysThrGlUArgCysLeuAspTrpGlUAlaGlY 157
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
419 GCCACAGCAGGTTGTCGATGATGGTGTGATGCTACACTCCTCAATCC 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
157 LyrProAspIleValAlaSPAspIleGlyAspAlaThrLeuLeuIleH 173
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
469 AAGGGCTTC.....GAATTCGAAACAGCGGCTGCTGTTCAGA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
174 GlUglYAlaLysAlaGlUglUglUglUglUglUglUglUglUgl 190
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
507 GCCAACAAGAGCTGACAACTCGAATACCGGCTTCTGCTACACTCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
190 pProGluSerThrAspAsnAlaGlUphelYsIleValleuThrIle 207
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
557 AGCAGGCTCTCAACCAAGCAAGCAACCACTGACACAGTGTGCTCGCG 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
207 rglAspIleLysAlaSPAspIleLysIleYrArgLysMetLysglUarg 223
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
607 ATGAACGGTGTTCGAGAGACAAACAAGAGTGTCCACCGCTCTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
224 LeuValGlYAlaSerGlUglUglUglUglUglUglUglUglUgl 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
657 GCTCGAGAGAGAGGCAACTCCTTCCAGCCATCAACGTCACAGCAG 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
240 mMetGlnGlUThrGlYAlaLeuLeuPheProAlaIleAsnValaSP 257
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
707 CTGTTCAAAGTCCAGTTCGATTAACATCTACGGCTGCGCACATCCCT 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
257 erValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerle 273
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
757 ATCGATGTATCAACGCTGCTTCGATGTCAATGCGCGGCAAGACAGC 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
274 ProAspIleLysMetArgAlaThrAspValMetIleAlaGlyLysVal 290
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
807 TCTGTCATCGGTTACGGCGATGTCGAGCAAGCGCTGCTCATCCCTCC 856

```

```

:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
290 aValValCysGlyTrpGlyAspValGlYLysGlyAlaAlaLeuL 307
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
857 GTGGCCAAAGCGCTGCGCTTATCATCACAGAACTCAGCCATCTGCGCT 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
307 ySGlnAlaGlYAlaArgValIleValThrGlUleAspProIleCysAla 323
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
907 CTCAGGCTGCCATGAAGGCTAACAGAGTCCCGCCGATCAGAGAGTGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
324 LeuGlnAlaLeuMetGlUglYleuGlnValleuProLeuGlnAspVal 340
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
957 CAAGATGTCCATATCTCTGTCATGTCACAGGAACTCGCATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1007 CTGTGACATGATGCGCCAGATGAAGATGAAGCTTATGCTGCTAATC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
340 lSerGlnAlaSPlePheValThrThrThrGlyAsnLysAspIleleu 357
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
357 etValAspHisMetArgLysMetLysAspAsnAlaIleValCysAsnIle 373
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1057 GGCCACTTCGATTAACGAAATGATACAGATGGCTGATGAATATCCAG 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
374 GlyHisPheAspAsnGlUleAspMetLeuGlUthrTrpProGl 390
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1107 CATCAGACATCCCAATCAACCCAGAAATACGACATGTGGAAATCCAG 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
390 YAlLysArgIleThrIleLysProGlInThrSPArgTrpValPhePro 407
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1157 ATGGCCAC...GCTATCCCTCTTGTGTCGAGGCGCGCTTCAACTT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
407 luthrAsnThrGlyIleIleValleuAlaGlUglYrArgLeuMetAsn 423
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1204 GCGTCGCTACAGAGTACCCCATCTTTCGTTATGTCATGTCATCACAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
424 GlyCysAlaThrGlnHisProSerPheValMetSerCysSerPheThr 440
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1254 CCAAGACATCGCTCAGCTCGACCTGTACGAAAAAGAA.....GGAATC 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
440 nglValIleAlaGlnLeuGlnLeuTrpLysGlyLysSerGlyLysT 457
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1298 TCGAAGAGAGGTTTACACACTTCGAAACATCTCGATGAAGAAGTCGT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
457 YrGlUglYLysValTrpValleuProLysHisLeuAspGlUysValAla 473
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1348 CGCCTCACCTCGGATCTCGATGTCCACCTTACAAAGCTTACACAGAA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
474 AlaleuHisLeuGlYLysLeuGlYAlaLysLeuThrLysLeuThr 490
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1398 GCAGGCTGATCAATCAACAGTTCGATGAAGGCTGCTTCAACTGATG 1447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
490 rGlAlaAspTrpYrIleSerValProIleGlUglYProTrpYrLysPro 507
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1448 CTACCGCTTAT 1458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
507 lSTyrArgTyr 510
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq\_name: /cgn2\_6/ptodata/2/pae/US60.COMB.pep:US-60-324-109-25353

seq\_documentation\_block:  
; Sequence 25353, Application US/60324109  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Stein, Joshua  
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-10(52726)B  
; CURRENT APPLICATION NUMBER: US/60/324,109  
; CURRENT FILING DATE: 2001-09-21  
; NUMBER OF SEQ ID NOS: 33196  
; SEQ ID NO 25353  
; LENGTH: 499

```

; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
US-60-324-109-25353

```

```

alignment_scores:
  Quality: 1520.00      Length: 487
  Ratio: 3.868          Gaps: 4
  Percent similarity: 80.698  Percent identity: 61.602

```

```

alignment_block:
US-09-759-990-1 x US-60-324-109-25353 ..

```

```

Align seg 1/1 to: US-60-324-109-25353 from: 1 to: 499

```

```

19 ACTGGTGCCTCATGTCAGTAATGCGCATCATCACTCCATGTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22 ThrSerSerGlyArgGluTyrLysValLysAspLeuSerGlnAlaAsp 38
CGGCGTAAGAGAACTTACCCTTGCTGAGAGAAATGCCAGGCTTTATG 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuMet 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTCTTCGTAGCGCTTATCCGCTTCTTAAGCATTTGAAGGCTGCAATC 168
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 IacYarGlaGluPheGlyProSerLysProPheAlaGlyAlaArgIle 71
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGTCCTCCATGACATGACAGTCACAGACGCGCTCATGAGACACT 218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThr 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTTGCTGATGTCAGATGGGCTCTGCAACATCTTCTCTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 ThrAlaLeuGlyAlaGluValArgTyrPysSerCysAsnIlePheSer 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAAGATACAGCCGCTGCTGCTATCGTGTGCGGCCCAACAGCACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
105 hrGlnAspHisAlaAlaAlaAlaAlaAlaArgAspSerAlaAla.... 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GAGAAAGCCAGCCGCTATCCAGCTTGCCTGGAAGGGCGAAACACTGCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 .....ValPheAlaTrpLysGlyLeuThrLeuGlu 129
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATTAAGTGGAGAACACATACCGCGCTCTACATGCCAGATGGTCAAC 418
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 uGluTyrTrpTyrPysThrGluArgCysLeuAspTrpGlyAlaGlyG 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 GCCCAGACAGAGTTGTCGATGATGGTGGTCACTGCTCACTCTCTATCC 468
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 LyrProAspLeuIleValAspAspGlyGlyAlaThrLeuLeuIleHis 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 AAGGCGTCT.....GAATGGAACAGCGCGGTGCTGTTCACGA 506
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 GluGlyValLysAlaGluGluAspTyrGlyLysThrGlyLysIlePro 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
507 GCCACAGAGAGTGCACAACTCGAATACCGCTGCTTCTGCTACACTCA 556
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 ProGluSerThrAspAsnAlaGluPheLysIleValLeuThrIleLeu 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
557 AGCAGGCTCTTCAACCAAGAACAGACACTGCGACACAGTCTGCCGCG 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 rGAspGlyLeuLysAlaAspProLysLysTyrArgLysMetLysGluArg 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
607 ATGAACGCTGTTCCGAGAGACAAACAGAGTGTCCACCGGCTCTACCA 656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyr 229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 GCTCGAAGAGAGGCAAACTCTCTCCAGCACAACGTCACAGCAAGAG 706
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
229 nMetGlnGluThrGlyAlaLeuLeuPheProAlaIleAsnValAsnAsp 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707 CTGTTCACAAGTCCAAAGTTCGATACATCTACGGCTGCCGCCACTCCCT 756
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

246 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757 ATCGATGATCAACCGTCTCCGATGTCATGATCGGCGGCAAGACAG 806
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807 TCTCGTCAATGAGTTTACGGCAGATGTCGGCAGAGGCTGCCCTCAAC 856
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
279 aValValCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaAlaLeu 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
857 GTGGCCAAAGGCGCTCGGCTTATCATCACAGAACTGACCAATCTGCG 906
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 yGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
907 CTCGAGCTCCATGGAAGGCTACAGGCTCCGCGCATGAGGAAGTCTG 956
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuProLeuGluAspVal 329
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
957 CAAGATGTCGATATCTTCTTACATGCACAGAAACTGCGATATCATCT 1006
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
329 lSerGlnAlaAspIlePheValThrThrThrGlyAsnLysAspIleLeu 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1007 CTGTTCACATGATGGCCAGATGAAGATTAAGCTATTGTCGTAACATC 1056
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1057 GGCCACTTCGATTAAGAAATGATACAGATGCGCTCATGAATATCCAG 1106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrPro 379
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1107 CATCAAGACACATCCCAATCAAGCCAGAAATACGACATGGGAATCCAG 1156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
379 yAllyLysArgIleThrIleLysProGlnThrAspArgTyrValPhePro 396
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1157 ATGGCCAC...CGTATCTCTCTTCTGTCGAGGCGCGCTTCTTAACCT 1203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
396 lThrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1204 GGCTCGCTACAGGTCAACCCATCTTCTGTTATGTCATATTCACAA 1253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
413 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThr 429
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1254 CCAGACACTGCTCAGCTCGACTCTTACGAAAGAGA.....GGAATC 1297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
429 nGluValIleAlaGluLeuGluLeuTrpLysGlyLysSerGlyLysT 446
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1298 TCGAAGAAGAGTTTACACACTTCCGAAGCATCTGATGAAGAAGTCGCT 1347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
446 yTrpLysLysValTyrValLeuProLysHisLeuAspGlyLysValAla 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1348 CGCTCCAGCTCGGATCTCTGATGTCACACTTACAAACCTTACACAG 1397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 AlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThrLys 479
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1398 GCAGGCTGACTATCAACGTTCCAGTTGAGGGTCCCTTACAAAGTCGAT 1447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
479 rGlnAlaAspTyrLysLeuValProIleGluGlyProLysProAla 496
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1448 CTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
496 lAsTyrArgTyr 499
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

seq_name: /cgn2_6/Plodata/2/paa/US60_COMB.pap:US-60-312-544-7754

```

```

seq_documentation_block:
; Sequence 7754, Application US/60312544

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Cao, Yongwei

```

```

; APPLICANT: Edgerton, Michael D

```

```

; APPLICANT: Hinkle, Gregory J.

```

```

; APPLICANT: Kovalic, David K.

```

```

; APPLICANT: Liu, Jingdong

```



seq\_documentation\_block:  
; Sequence 26717, Application US/60324109  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jindong  
; APPLICANT: Stein, Joshua  
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-10(52726)B  
; CURRENT APPLICATION NUMBER: US/60/324,109  
; CURRENT FILING DATE: 2001-09-21  
; NUMBER OF SEQ. ID NOS: 33196  
; SEQ. ID NO: 26717  
; LENGTH: 504  
; TYPE: PRN  
; ORGANISM: Zea mays  
; FEATURE:  
US-60-324-109-26717

alignment\_scores:  
Quality: 1520.00 Length: 487  
Ratio: 3.868 Gaps: 4  
Percent Similarity: 80.698 Percent Identity: 61.602

alignment\_block:  
US-09-759-990-1 x US-60-324-109-26717 ..

Align seg 1/1 to: US-60-324-109-26717 from: 1 to: 504

```
19 ACTGGTCCTCCATTCAGATGACAGATGCGCAGATCCATCCATGCTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
27 ThrSerSerGlyArgGlnTyrLysValLysAspLeuSerGlnAlaAspH 43
CGCCGCTAGAGACTTACCTGCTGAGAAAGATGCGCAGGCTTATG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43 eGlyArgLeuGlnIleGlnLeuAlaGlnValGlnMetProGlyLeuMet 60
TTCTCTGAGAGCGTTATTCGCGCTTCTTGAAGCATGAGGCTGACATC 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 LacysArgAlaGlnPheGlyProSerLysProPheAlaGlnAlaArg 76
TCTGTCCTCCATTCAGATGACAGATGCGCAGATCCATCCATGCTTCT 218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
77 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGlnThr 93
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTGTCGTGATGTCAGATGAGGCTTCTGCAACATCTTCTCTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 uThrAlaLeuGlnIleGlnValAlaArgTyrPysSerCysAsnIlePhe 110
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAGATACAGCGCGCTGCTGCTATCGTTGCGGCCACAGGACACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 hrcInAspHisAlaAlaAlaIleAlaAlaArgAspSerAlaAla.... 124
GAGAGAGCAGCGGATATCCAGTCTGCTGAGAGGCGAAGACACTCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 .....ValPheAlaTyrPylGlnThrLeuGln 134
AGATATCTGGAGACATACCGCGCTCTCAGATGCGCAGATGCTCAAG 418
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 uGlnTyrTrpTyrPylGlnArgCysLeuAspTyrPylGlnAlaGln 151
GCGCAGCAGAGGTCGTGATGATGAGTGTGATGATGATGATGATGATG 468
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 LysProAspLeuIleValAlaAspAspGlyAlaSerAlaThrLeuLeu 167
AAGGAGCTTC.....GAATTCGAAGACGCGGCTGCTTCTCAGA 506
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
168 GlnGlnValLysAlaGlnGlnAspTyrGlnTyrGlnTyrGlnTyrGln 184
```

```
507 GCCAAGAGAGCTGACAACTCGAATACCGCTGCTTCTGCTACACTCA 556
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 ProGlnSerThrAspAsnAlaGlnPheLysIleValLeuThrIleLeu 201
AGCAGGCTTTCACCAAGACAGAACATCGGCGCAGATGCTGCGCGC 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 rAspGlyLeuLysAlaAspProLysLysTyrArgLysMetLysGlnArg 217
ATGAACGGTGTTCGAGAGACAGAACAGAGTTCACCGCTCTACCA 656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 LeuValGlnValSerGlnTyrThrThrThrGlyValLysArgLeuTyrG 234
GCTCGAAGAGAGGCGCAACTCTCTTCCAGCAGCATCAAGCTCAAGCAG 706
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 nMetGlnTyrThrGlyAlaLeuPheProAlaIleAsnValAsnAsp 251
CTGTATCAAGATTCAGATTCATCAATCTACGCTGCGCGCATCTCCCT 756
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 eValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 267
ATCAGTGTATCAACCGTCTCGATGTCATGATGCGCGCAGACAGC 806
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 ProAspGlyLeuMetAlaThrAlaThrAspAlaMetIleAlaGlyValAl 284
TCTCGTCATAGGCTTACGCGCATGCGCAAGGCGCTGCAATCCCTCC 856
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 aValValLysGlyTyrGlyAspValGlyLysGlyAlaAlaAlaLeu 301
GTGGCCAGAGCGCTCGCTATCATCAAGAACTGACCCATCTGCGCT 906
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 yGlnAlaGlyAlaArgValIleValIleThrGlnIleAspProIleCysAla 317
CTCCAGGCTGCGCATGAGAGCTTACAGTCCGCGCATCGAGAGACTG 956
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
318 LeuGlnAlaLeuMetGlnGlyLeuGlnValLeuProLeuGlnAspValVa 334
CAAGATGTCGATATCTTCTGTTACATGACAGAACTCGATATCATCT 1006
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
334 lSerGlnAlaAspIlePheValIleThrThrGlnLysnLysAspIleLeu 351
CTGTGATCATGATGCGCCAGATGAGAAATGAGATGATGCTGCTGATCA 1056
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 eValAspHisMetAlaGlyLysMetLysAsnAsnAlaIleValCysAsn 367
GGCCTACTGATTAAGAAATGATACAGATGCGCTCATGAAATACCAG 1106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
368 GlyHisPheAspAsnGlnIleAspMetLeuGlyLeuGlnThrTyrProG 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1107 CATCAAGCAGATCCCAATCAAGCCAGAAATACGATGAGGAATTCACG 1156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1157 ATGGCCAC...GCTATCTCTCTTCTGCTGAGGCGCTTCTTAACCTT 1203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 lThrAsnThrGlyIleIleValLeuAlaGlnGlyArgLeuMetAsnLeu 417
GGTCCGCTACAGGTCACCAATCTTCTGATGTCATGATGATGATGAT 1253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 GlyCysAlaThrGlnHisProSerPheValMetSerCysSerPheHis 434
CCAGACACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1254 nGlnValIleAlaGlnLeuGlnLeuTyrPylGlnLysSerSerGlyLys 451
TCGAGAGAGAGGTTACACACTTCGAGACATCTCGATGAGAGAGTGGCT 1347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
451 yGlnLysValValTyrValLeuProLysHisLeuAspGlnLysValAla 467
CGCTCCAGCTGCGATCTCTGATGTCACCTTACAAAGGTTACAGAGA 1397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
468 AlaLeuHisLeuGlnLysLeuGlnAlaLysLeuThrLysLeuThrLys 484
GCAGGCTGACTACATCAACGTTCCAGTGAAGGCTCTTACAGTCTGATG 1447
```

```

|||||
484 rginlaasptylleServalProtleGluGlyProTylLysProAlaH 501
1448 CTACCGTTAT 1458
|||||
501 lStYrArGTyr 504

```

```
seq_name: /cgn2_6/prodata/2/paa/US60.COMB.pep:US-60-324-109-32880
```

```
seq_documentation_block:
```

```
; Sequence 32880, Application US/60324109
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Cao, Yongwei
```

```
; APPLICANT: Edgerton, Michael D
```

```
; APPLICANT: Hinkle, Gregory J.
```

```
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Liu, Jindong
```

```
; APPLICANT: Stein, Joshua
```

```
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
```

```
; FILE REFERENCE: 38-10(52726)B
```

```
; CURRENT APPLICATION NUMBER: US/60/324,109
```

```
; NUMBER OF SEQ ID NOS: 33196
```

```
; SEQ ID NO 32880
```

```
; LENGTH: 504
```

```
; TYPE: PRT
```

```
; ORGANISM: Zea mays
```

```
; FEATURE:
```

```
US-60-324-109-32880
```

```
alignment_scores:
```

```
Quality: 1520.00
```

```
Ratio: 3.868
```

```
Percent Similarity: 80.698
```

```
Percent Identity: 61.602
```

```
Length: 487
```

```
Gaps: 4
```

```
Alignment Identity: 61.602
```

```
Alignment Length: 487
```

```
Alignment Ratio: 3.868
```

```
Alignment Percent Similarity: 80.698
```

```
Alignment Percent Identity: 61.602
```

```
Alignment Length: 487
```

```
Alignment Gaps: 4
```

```
Alignment Alignment Identity: 61.602
```

```
Alignment Alignment Length: 487
```

```
Alignment Alignment Ratio: 3.868
```

```
Alignment Alignment Percent Similarity: 80.698
```

```
Alignment Alignment Percent Identity: 61.602
```

```
Alignment Alignment Length: 487
```

```
Alignment Alignment Gaps: 4
```

```
Alignment Alignment Alignment Identity: 61.602
```

```
Alignment Alignment Alignment Length: 487
```

```
Alignment Alignment Alignment Ratio: 3.868
```

```
Alignment Alignment Alignment Percent Similarity: 80.698
```

```
Alignment Alignment Alignment Percent Identity: 61.602
```

```
Alignment Alignment Alignment Length: 487
```

```
Alignment Alignment Alignment Gaps: 4
```

```
Alignment Alignment Alignment Alignment Identity: 61.602
```

```
Alignment Alignment Alignment Alignment Length: 487
```

```
Alignment Alignment Alignment Alignment Ratio: 3.868
```

```
Alignment Alignment Alignment Alignment Percent Similarity: 80.698
```

```
Alignment Alignment Alignment Alignment Percent Identity: 61.602
```

```
Alignment Alignment Alignment Alignment Length: 487
```

```
Alignment Alignment Alignment Alignment Gaps: 4
```

```
Alignment Alignment Alignment Alignment Alignment Identity: 61.602
```

```
Alignment Alignment Alignment Alignment Alignment Length: 487
```

```
Alignment Alignment Alignment Alignment Alignment Ratio: 3.868
```

```
Alignment Alignment Alignment Alignment Alignment Percent Similarity: 80.698
```

```
Alignment Alignment Alignment Alignment Alignment Percent Identity: 61.602
```

```
Alignment Alignment Alignment Alignment Alignment Length: 487
```

```
Alignment Alignment Alignment Alignment Alignment Gaps: 4
```

```

Align seg 1/1 to: US-60-324-109-32880 from: 1 to: 504

```

```

19 ACTGGTGTCTTCATTCAGTACAGTAATGCCGACATCAACCTCCATGTTCT 68
27 TTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 43
69 CGCCGCTAAGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118
43 eglYalgluGluIuIleIuLeuAlaGluValGluMetProGluLeuMet 60
119 TTTCTGTGAGCGTTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 168
60 lAcysArGAlaGluPheGlyProSerIysProPheAlaGlyAlaArgIle 76
169 TCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 218
77 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrIle 93
219 CACAGCTCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 268
93 uThrAlaLeuGlyAlaGluValAlaArgItrPcysSerCysAsnIlePheSer 110
269 CACAAGATACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
110 hGlnAAspHisAlaAlaAlaIleAlaIleAlaArgAspSerAlaAla..... 124
319 GAGAAGCCAGCCGGTATCCAGTCTTCGCTGGAAGGGCGAAGACATCCGCC 368
125 .....ValPheAlaItrPcysGlyGluThrLeuG 134
369 AGAATAGTGGAGAACATACGAGCTCTCCACATGCGCCAGATGCTGACAG 418
134 uGluItrPcysThrGluArgCysLeuAspItrPcysGlyGluAlaGly 151

```

```

419 GCCACAGCAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 468
151 lYProAspLeuIleValAspAspPcylGlyAspAlaItrLeuIleHis 167
469 AAGGCTTC.....GAATTCGAACACAGCCGCTGCTGCTGCTGCTGCTGCT 506
168 GluGlyValLysAlaGluGluAspTyrGluLysThrGlyLysIleProAs 184
507 GCCACAGCAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 556
184 ProGluSerThrAspAsnAlaGluPheLysIleValLeuThrIleIle 201
557 AGCAGGCTTCAACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAG 606
201 rGAspGlyLeuLysAlaAspProLysItrPcysGlyMetLysGluArg 217
607 ATGAAGCTGTTTCGAGAGACACACACACACACACACACACACACACACACAC 656
218 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuItrG 234
657 GCTGAGAGAGAGGCAAACTCTCTCTCCAGCCTCAACAGCTCAACAGCAGC 706
234 nMetGlnItrGlyAlaLeuLeuPheProAlaIleAsnValAsnAsp 251
707 CTGTTCAAAAGTCAAGTTCGATTAACATCTACGCTGCGCCACTGCTT 756
251 eValThrLysSerLysPheAspAsnLeuItrGlySerAlaGlnSerLeu 267
757 ATCGATGTTATCAACGCTCTCCGATTCATCATCAGAACTCGAACCAATCGCT 806
268 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 284
807 TCTGCTCATGGGTACGGGATGTCGCAAGGCTGCGCTGCTGCTGCTGCTGCT 856
284 aValValCysGlyTyrGlyAspValGlyLysGlyAlaAlaLeuL 301
857 GTGGCCAGGCGCTCGCTTATCATCATCAGAACTCGAACCAATCGCT 906
301 ySglnAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 317
907 CTCAGGCTGCTGATGAGGCTACGAGCTGCGCCGCTGCTGCTGCTGCTGCT 956
318 LeuGlnAlaLeuMetGluGluLeuGlnValLeuProLeuGlnAspValVa 334
957 CAAGATGTCGATATCTTCTTACATGACAGAACTCGAATATCATCT 1006
334 lSerGluAlaAspIlePheValThrThrThrLysAsnIleIle 351
1007 CTGTTCATCATGATGCGCCAGATGAAGATTAAGCTATGTCGATATCATCT 1056
351 eValAlaAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 367
1057 GGCACCTTCGATTAAGAAATGATACAGATGCGCTGCTGCTGCTGCTGCTGCT 1106
368 GlYHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrThrProAl 384
1107 CATCAAGCAGATCCCAATCAAGCAGAAATACAGATGCTGGAATCCAG 1156
384 yValLysArgIleThrIleLysProGlnThrAspArgItrPcysAlaPheProG 401
1157 ATGGCCAC...GCTATCTCTCTCTGCTGAGGCGCGCTTCTTAACCTT 1203
401 lUthrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 417
1204 GGCTGCGCTACAGGTCACCACTTCTGTTATGTCATATGATATGATATGAT 1253
418 GlyCysAlaIleThrGlyHisProSerPheValMetSerCysSerPheThrAs 434
1254 CCAGACATCTGCTCAGCTCAACCTCAAGAAAGAAAGAA.....GGAATTC 1297
434 nGlnValIleAlaGlnIleGluIleuLeuItrPcysGluLysSerGlyLys 451
1298 TCGAGAAAGAGTTTACACACTTCGGAAGCATCTTCGATGAAGAAAGTCCGT 1347

```

```

451 yrcgluylslsvaltyrvalleuprollysleuaspiglulysvalala 467
1348 CGCCTCCACCTCGGATCTCTGATGTCCACCTTACAAAGCTTACACAGAA 1397
468 Alaleuhtislenuglylsleuglyalalysleuthrllysleuthrls 484
1398 GCAGGCGTACTACATCAACGCTCCAGTGTGAGGCTCTTACAAAGCTGATG 1447
484 rglnalaaaprylleserlevalproleuglulglyprolyrlyspaalaa 501
1448 CTACCGCTTAT 1458
501 lstryarqlyr 504

```

seq.name: /cgn2.6/ptodata/2/paa/US097\_COMB.pep:US-09-708-427-22305

```

seq_documentation_block:
; Sequence 22305, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22305
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..467
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..467
; OTHER INFORMATION: Ceres Seq. ID 1840723
; US-09-708-427-22305

```

alignment\_scores:

Quality:	1519.00	Length:	470
Ratio:	3.966	Gaps:	4
Percent Similarity:	81.489	Percent Identity:	63.404

alignment\_block:

US-09-759-990-1 x US-09-708-427-22305 ..

```

Align seg 1/1 to: US-09-708-427-22305 from: 1 to: 467

70 GACCGTAAGGAAGTACCTGCTGAGAGAAATGCGAGTCTTATGCT 119
7 GAGTGTGAGGCTTATCCGCTTCTAAGCCATTGAAGGCTGAGAAATCT 169
120 TCTTGTGAGGCTTATCCGCTTCTAAGCCATTGAAGGCTGAGAAATCT 169
23 acysargthrighupheglyproserglnprophelyslyalaaqglyler 40
170 CTGGTCCCTCCATGACATGACATGACAGAGCCGCTCTCATGAGACACTG 219
40 hrglyserleuhtismetlhrileglnthralvalleuileglnthrlreu 56
220 ACAGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 269
57 Thratalaenuglyalagluvalargtrpyssercysasnillepserth 73
270 ACAAGATACAGCGCGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 319
73 rglnasprhsalaaalaaalaaalaaalaaalaaalaaalaaalaaalaa 87
320 AGAAGCCAGCGCGTATCCAGATCTTGGCTTGAGAGGCGAAACACTCCCA 369

```

```

88 .....Valphealetripysglulthrlleugln 97
370 GAATACHTGGGAGACACATACCGCGCTGTACATGGCCACAGTGTCAAG 419
98 Glutrytrprrpysrthrighupheglyalalasprrprrprrprrprrprr 114
420 CCCACAGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 469
114 yproaspheulelevalaspaspiglylaspalatrhlleuilehtisg 131
470 AGGCTTCGAA.....TTGAAACAGCCGCTGCTTCACAG 507
131 lucglyvalalagluglulilepheglulysleuthrlglnvalproasp 147
508 CCAACAGACAGTGCACACCTGATATACCGCTGCTTCCTGATCACTCA 557
148 Prothrserthrhrasprnproglupheglnlilevalleuhtisrleilely 164
558 GCAGGTCTTCAACCAAGACAGACACACACAGTGTGCTGCGCGCA 607
164 sguluglyleuglnvalasprrprrprrprrprrprrprrprrprrprrprr 181
608 TGAACGCTGTTCGAGAGACACACACAGTGTGCTGCGCGCTTACAG 657
181 euvalglyvalsergluglulthrlthrighupheglyalalasprrprrprr 197
658 CTCGAGAGAGGAGGCAAACTCTCTCCAGCCATTAAGTCAACAGAGC 707
198 Metglnlnasnuglythrleuupheprrprrprrprrprrprrprrprrprr 214
708 TGTTCACAAAGTTCAGTTCGATATACATGATGATGATGATGATGATGATGAT 757
214 rValthrlysserlysrpheaasnleuthrlglysarhnsrleup 231
758 TCGATGTATCAACCGTCTCCGATGTATGATGATGATGATGATGATGATGAT 807
231 rCaspiglyleuhtismetlhrileglnthralvalleuileglnthrllysvalala 247
808 CTCGTATGAGGTTACGGGATGTCGCAAGGCTGCGCTCAATCCCTCG 857
248 Valillecysgllyrlyaspvalalglylselcysalialalalamelely 264
858 TGGCCAGGCGCTGCGCTTATCATCACAGAACTGCACCCAAATCTGCGCTC 907
264 sthrilaglialargvalillevallthrghlileasprollecysalil 281
908 TCCAGGCTGCCATGGAAGGCTACACAGTCCCGCGCATGAGACAGTCTC 957
281 euqlnalaueuhtismetlhrileglnlileuthrlleuglnaspvalal 297
958 AAGGATGTGATATCTGTTACATGACAGAAATGCGATATGATCTC 1007
298 Serghlualasprrlephevalthrthrighupheglyalalasprrlelle 314
1008 TGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1057
314 tValasprhsismetlhrileglnthralvalleuileglnthrllysvalala 331
1058 GCCACTTCGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
331 lYhnsprhsasnnglulileasprmetleuglyleuglulthrltyrprogly 347
1108 ATCAAGACATCCCAATCAAGCAAGATACAGATGATGATGATGATGATGATGAT 1157
348 Vallysarglilethrileusproglulthrasprargtrpvalpheprool 364
1158 TGGCAAGCT...ATCCTCTTCTGCTGAGGCGCGCTTCTTAACTTGG 1204
364 uthrlysalaglylilevalleuhtismetlhrileglnthrlleuhtisnleug 381
1205 GCTGCGTACAGATCAACCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1254

```

```

381 lYcysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAsn 397
1255 CAGACACTGCTCAGCTCGACTCTACGAAAGAGA.....GGAATCT 1298
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 GluValIleIleAlaGlnLeuGlnLeuTyrPasnGlnLysAlaSerGlyLysTyr 414
1299 CGAGAAAGAGGTTTACACACTTCCGAAGCATCTTCGATGAAGAGTCCGCTC 1348
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 rGluLysLysValTyrValLeuProLysHisLeuAspLysValAlaIle 431
1349 GCCTCAGCTCGATCTCTCGATGTCACCTTACAAAGCTTACACAGAG 1398
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 euLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAsp 447
1399 CAGGCTGACTACATCAAGCTTCCAGTTGAGGCTCTTACAAAGTCCGATGC 1448
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 GlnSerAspLysValSerIleProIleGlnGlyProTyrLysProProH 464
1449 TTACGCTTAT 1458
|||||:|||||:
464 sTyraTyrTyr 467

seq_name: /cgn2_6/ptodata/2/paa/US098_COMB pep:US-09-855-768-748

seq_documentation_block:
: Sequence 748. Application US/09855768
: GENERAL INFORMATION:
: APPLICANT: Hauge, Brian M.
: APPLICANT: Helne, Danielle L.
: APPLICANT: Lin, Jie-Yi
: APPLICANT: Parsons, Jeremy D.
: APPLICANT: Wang, Ming Li
: TITLE OF INVENTION: PLANT GENOME SEQUENCES AND USES THEREOF
: FILE REFERENCE: 38-10(52047)B
: CURRENT FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 09/760,427
: PRIOR FILING DATE: 2001-01-13
: NUMBER OF SEQ ID NOS: 1107
: SEQ ID NO 748
: LENGTH: 485
: TYPE: PRT
: ORGANISM: Glycine max
: US-09-855-768-748

alignment_scores:
Quality: 1515.00 Length: 482
Ratio: 3.855 Gaps: 4
Percent Similarity: 81.535 Percent Identity: 61.826

alignment_block:
US-09-759-990-1 x US-09-855-768-748 ..
Align seg 1/1 to: US-09-855-768-748 from: 1 to: 485

34 GAGTACAGAAATTCGCGACATCACTCATGTCTCGCGCGTGAAGAACT 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GluTyrLysValLysAspLeuSerGlnAlaAspPheGlyArgLeuGlnI 29
84 TACCTTGCTGAGAAAGTCCAGGCTTATGTTCTTCGTTGCTGAGAGCT 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 euLLeuAlaGlnValGlnMetProGlyLeuMetAlaCysArgThrGln 46
134 ATTCGCTTCTTAAGCCATTTGAAGGCTGTCAGATCTCTGTTCCCTCCAC 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 heGlyProSerGlnProPheLysGlyAlaArgIleThrGlySerLeuHis 62
184 ATGACAGTCCAGACAGCCGCTCATCGAGACATCAGACAGCTTGTGTCG 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 MetThrIleGlnThrAlaValLeuIleGlnIleThrLeuThrAlaLeuGly 79
234 TGAGTGTAGATGGGCTTCCTGCAACATCTTCTTACAAAGATACAGCCG 283

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 aGluValArgTyrPysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATTCGTTGTGGCCCAACAGGACACACAGAGAACGCCACCG 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 IaAlaIleIleAlaArgAspSerAlaIa..... 105
334 ATCCAGCTCTCGCTGGAGAGGCGAAGACATCCAGATATCTGGAGAA 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTyrPylGlyThrLeuGlnGlyTyrTyrPyr 120
384 CACATACCGGCTCTCATATGCGCAGATGTCAGGCCACAGCAGAGTTG 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 sThrGlnArgAlaLeuAspTyrGlyProGlyGlyLeuProAspLeuIle 137
434 TCGATGATGGGTGATGTCACATCTTCATCTCCAAAGGCTTCGAA... 480
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 AlAspAspGlyGlyAspAlaThrLeuLeuIleHisGlnGlyValLysAla 153
481 .....TTCGAAACAGCCGCTGTTCAGAGCCACAGAGAGCTGA 521
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 GluGluLeuTyrGlnLysThrGlyLeuProAspProAsnSerThrAs 170
522 CAACTCGAATACCGCTGCTTCTTGTCTACATCAAGAGGTCTTCAAC 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 rAsnAlaGlnPheGlnIleValLeuThrIleIleArgAspGlyLeuLys 187
572 AAGACAGAACACACACGCGCAGATGCTCCGCGCATGAACGCTGTTCC 621
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 hAspProThrArgTyrArgLysMetLysGlnArgValGlyValSer 203
622 GAGAGACACACACAGGTGTCACGCGCTCTACAGCTGAGAGAGAGG 671
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 GluGluThrThrThrGlyValLysArgLeuTyrGlnMetIleAsnGln 220
672 CAACTCTCTTCCAGCATCAAGCTCAAGAGCTGTATCAAGATGCA 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 YThrLeuLeuPheProIleAlaHisValAsnAspSerValThrLysSer 237
722 AGTTCGATTAACATTCAGGCTGCGCGCTCCCTTATCGATGTATCAAC 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 ySPheAspAsnLeuTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253
772 CGTGTTCGATGATGATGATGCGGCGCAGAGAGCTGTGATGAGTTA 821
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValAlaValAla 270
822 CGGCGATGTGCGCAAGGCTGCGCTCATATCCCTCGTGCGCAAGCGCTC 871
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 rGlyAspValGlyLysGlyCysAlaIleAlaIleMetLysGlnAlaGly 287
872 GCGTTATCATCACAGACTCGACCCCAATCTGCGCTTCCAGCGCTGCCATG 921
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 rGValIleValIleThrGlnLeuAspProIleCysAlaLeuGlnAlaLeuMet 303
922 GAAAGCTACAGAGTCCGCGCATCGAGAGAGTGTCAAGATGTCATAT 971
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 GluGlyLeuGlnValLeuThrLeuGlnLysAspValValSerGlnAlaAsp 320
972 CTTCGTTATCATCACAGAAACTCGCATATCATCTCTGTGACATGATGG 1021
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ePheValThrThrThrGlyAsnLysAspIleIleMetValAspHisMet 337
1022 CCCAGATGAAGATTAAGGCTATGTGGGTGATCAATCGGCGCACTTGATAC 1071
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 rGlySerLysAsnAlaIleValAlaCysAsnIleGlyHisPheAspAsn 353
1072 GAAATGTATACAGATGCGCTCATGAAATATACAGAGCATCAAGCATGCC 1121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 GluIleAspMetLeuGlyLeuGlnLysAsnTyrProGlyValLysArgIle 370
1122 AATCAAGCCAGATTCAGCATGTGGGAATTCGCCAGATGCCAC...GCTA 1168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```





```

354 GluIleaspMetLeuGlyLeuGluAsnTyrProGlyValIlysArgIleTh 370
1122 AATCAACGCGAATATACGATGATGGAAATCCAGATGGCCAC...GCTA 1168
|||||
370 rIleLysProGlnThrAspArgTTPValPheProGlnThrAsnThGly 387
1169 TCCCTCTTCTGCTGAGGCGCCCTTCTTAACCTGGCTGCGCTACAGGT 1218
|||||
387 leIleValIleuAlaGluGlyArgIleuMetAsnLeuGlyCysAlaThrGly 403
1219 CACCAATCTTCTGTTATGTCATTCATTCACAAACACAGACACTGCTCA 1268
|||||
404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaG 420
1269 GCTGACCTTACGAAAGAGA...GGAAATCTCGAAGAGAGATT 1312
|||||
420 AluGlnLeuTrrPlySgluLysSerThGlyLysTyrGluLysLysVal 437
1313 ACACACTTCCGAGATCTCGATGAAGAAGTCGCTCCGCTCCACCTCGGA 1362
|||||
437 yValIleuProLysHisLeuAspGluLysValAlaIleuHisLeuGly 453
1363 TCTCTCGATGTCACACTTACAAAGCTTACAGAGAGAGCTGACTACAT 1412
|||||
454 LysLeuGlyAlaLysLeuThrGlnLeuSerLysSerGlnAlaAspTyr 470
1413 CAACGTTCCAGTGTGAGGCTCTTACAGCTGATGCTTACCTTAT 1458
|||||
470 eSerValProValGluGlyProTyrLysProAlaHisTyrArgTyr 485

seq_name: /cgn2.6/prodata/2/paa/US097_COMB.pep:US-09-708-427-22306

seq_documentation_block:
; Sequence 22306, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22306
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..450
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..450
; OTHER INFORMATION: Ceres Seq. ID 1840724
US-09-708-427-22306

alignment_scores:
Quality: 1486.00 Length: 459
Ratio: 3.963 Gaps: 4
Percent Similarity: 81.699 Percent Identity: 63.181

alignment_block:
US-09-759-990-1 x US-09-708-427-22306 ..
Align seg 1/1 to: US-09-708-427-22306 from: 1 to: 450

103 ATGCCAGGTCTTATGCTTCTGAGCGTATTCGCTTAAAGCAT 152
|||||
1 MetProGlyLeuMetAlaCysArgThrGluPheGlyProSerGlnProPh 17
153 GAAGGTTCCGAATCTGCTGCTCCCTCCACATGACATGACAGTCCAGCGC 202
|||||

```

```

17 eLysGlyAlaArgIleThrGlySerLeuHisMetThrIleGlnThrAlaV 34
203 TCCCTATCGACAGACTCACAGCTCTGCTGCTATGTCAGATGGGCTTCC 252
|||||
34 AlIleuIleGluThrLeuThrAlaLeuGlyAlaGluValAlaGlyTrrCys 50
253 TGCACATCTTCTTACACAGATACAGCCGCTGCTGCTATGCTTGC 302
|||||
51 CysAsnIlePheSerThrGlnAspHisAlaAlaIleAlaIleAlaArgAs 67
303 CCCAACAGGACACACAGAGAGCCAGCCGCTATCCCGTCTTGCCTGGA 352
|||||
67 pSerIleAlaA.....ValIleAlaThrPl 75
353 AGGCGAAACACTCCAGATATGAGAGAACATACCCGCTGCTACA 402
|||||
75 ySglGlyThrLeuGlnGluTyrTrrPrrCysThrGluAlaGluAlaLeuAsp 91
403 TGGCCAGATGTCAAAGGCCACAGAGGTTGTCGATGATGGTGATGC 452
|||||
92 TrrGlyProGlyGlyProAspLeuIleValAspAspGlyAspAl 108
453 TACACTCTCATCTCCAAAGGCTTCGA.....TTGCAACAG 490
|||||
108 aThrLeuLeuIleHisGluGlyValIlyAlaGluGluIlePheGluTrr 125
491 CCGGTGCTGTCCAGAGCCACAGAGCTGACACACTCGAATACCGCTGC 540
|||||
125 hGlyGlnValProAspProThrSerThrAspAsnProGluPheGlnIle 141
541 GTTCTTGTACACTCAAGAGCTTCTTACCAACAGACAGAACCACTGGCA 590
|||||
142 ValLeuSerIleIleLysGluGlyLeuGlnValAspProLysTyrYr 158
591 CACAGTGTGCGCGCATGAAAGGTTGTCGAAGACACACACAGCAGTGC 640
|||||
158 sLysMetLysGluArgLeuValGlyValSerGluGluThrThrGly 175
641 TCCACCGCTCTTACAGCTCGAGAGAGGAGGAACTCTTCTCCAGCC 690
|||||
175 AlLysArgLeuTyrGlnMetGlnGlnAsnGlyThrLeuLeuPheProAla 191
691 ATCAACGTACAGACGCTGTTACAAAGTCCAAAGTTCATATCACTACGG 740
|||||
192 IleAsnValAsnAspSerValThrLysSerLysPheAspAsnLeuTyr 208
741 CTGCGGCCACTCCCTTATGATGATCAACCGGCTTCGATGATCAATGA 790
|||||
208 yCysArgHisSerLeuProAspLysLeuMetArgAlaThrAspAlaMet 225
791 TCGGGGCGCAAGACAGCTCGTCATGGGTTACGGCGATGTCGGCAAGGC 840
|||||
225 leAlaGlyLysValAlaValIleCysGlyLysGlyAspValGlyGly 241
841 TGGCGTCATCCCTCCGTGGCCAAAGGCTCGGCTTATCATACAGAACT 890
|||||
242 CysAlaAlaAlaMetLysThrAlaGlyAlaArgValIleValThrGlu 258
891 CGACCAATCTGGGCTCTCCAGGCTGCGCATGGAAGCTTACAGGCTCGCC 940
|||||
258 eAspProIleCysAlaLeuGlnAlaLeuMetGluGlyLeuGlnValLeu 275
941 GCATCGAGAGAGTGTCAAGGATGTCATATCTTCTTACATGCAAGGA 990
|||||
275 hLeuGlnAspValValSerGluAlaAspIlePheValThrThrThrGly 291
991 AACTGGATATATATCTCTGTTCATGATGAGCCAGATGAAGATTAAGGC 1040
|||||
292 AsnLysAspIleIleMetValAspHisMetArgLysMetLysAsnAspAl 308
1041 TATTGCGGTAACTACGCGCCTTGCATGCAATGATGATACAGATGCC 1090
|||||
308 AlIleValCysAsnIleGlyHisPheAspAsnGluIleAspMetLeuGlyL 325

```



```

1041 TATTGCGTACATCGGCACTTGATTAAGAAATGATACAGATGGCC 1090
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
308 attlevalcyasnilegylhisphenaspanglileasmetleugly 325
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1091 TCATGAATCCCGGATCAAGCAATCCCAATTAAGCCAAATACGC 1140
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
325 eucluthtyrproglvalylsargilethrllelsproglinthrasp 341
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1141 ATGTGGATATCCAGATGGCCAC...GCTATCTCTCTTGTGTGAGGG 1187
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
342 Argtrvalalheproglutthrasnthrclylellevalleuallaglucl 358
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1188 CGGCTTTTAACTTGGCTGCTGCTACAGGTACCCATCTTTCGTATGT 1237
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
358 yalglemetalsnleuglycysalathrglyhisproserphenalmetls 375
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1238 CAATGTCATCCAAACCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCT 1287
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 ercysserphenethrasnlevalleallaglnleugluleutrpysglu 391
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1288 AGA....GGAATCTCGAGAGAGAGGTTTACACACTTCCAGACATCT 1331
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
392 LysSerSerGlyTyrGlyLysValTyrValleuProLysHisle 408
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1332 CGATGAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1381
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 uasrglulylvalalalaleuileuglyllysleuallalalysleu 425
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1382 CAAGCTTACAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1431
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
425 hrlyserleuthrlyserleuallasprryllelservalprolleugly 441
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
1432 CCTTCAAGTCTGATGCTTACCGTTAT 1458
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
442 ProTyrLysProAlaHisTyrArgTyr 450
      ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq.name: /cgn2.6/ptodata/2/paa/US096\_COMB.pep:US-09-620-394B-3583

seq\_documentation\_block:

```

: Sequence 3583, Application US/09620394B
: GENERAL INFORMATION:
: APPLICANT: ALEXANDROV, Nikolai
: TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
: FILE REFERENCE: 2750-1067P
: CURRENT APPLICATION NUMBER: US/09/620,394B
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 9131
: SEQ ID NO 3583
: LENGTH: 467
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..467
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..467
: OTHER INFORMATION: Ceres Seq. ID 1387217
US-09-620-394B-3583

```

alignment\_scores:

```

Quality: 1457.00      Length: 466
Ratio: 3.865          Gaps: 4
Percent Similarity: 80.901      Percent Identity: 62.232

```

alignment\_block:

```

US-09-759-990-1 x US-09-620-394B-3583
Align seg 1/1 to: US-09-620-394B-3583 from: 1 to: 467

```

```

19 ACTGCTCTCCATTCAGATACAGAAATGCCAGATCAACCTCCATGTTCT 68
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerAlaAsp 24
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
69 CGGCGCTAAGCAATTAACCTTGTCTGAGAAAGAAATGCCAGCTTATG 118
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
24 eglYarGleuGluLeuGluLeuAlaGluValGluMetProGlyLeuMet 41
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 TTTCTTCGAGAGGTTATTCCTGCTTACAGCAATTCAGAGGTTGTCAGATC 168
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 lacysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgle 57
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 TCTGGTCTCTCCATCAATGACATCCAGACAGCCGCTCTCATGACAGACT 218
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
58 ThrGlySerLeuHisMetThrIleGlnIleThrAlaValleuIleGluThrle 74
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 CACAGCTCTGAGTGTGATGTCAGATGGGCTTCCCTCCACATCTTCTCTA 268
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 uThrAlaLeuGlyAlaGluValArgTyrPysSerCysAsnIlePheSer 91
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 CACAAGATACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 hrGlnAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla.... 105
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 GAGAAAGCCAGCCGGTATCCCACTCTGCTGCTGAGAGGCCAGAACTCTCC 368
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 .....ValPheAlaIleTrpLysGlyLysThrleuGlu 115
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 AGAATACGTGGAGAGACATACCGGCTCTACATGCGCAGATGTCAG 418
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 nglutyrtrprrpcysthrGlnArgAlaLeuAsprrpGlyProGlyGly 132
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 GCCACACAGAGGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 468
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 LyrProAspLeuIleValAspAspArgLysGlyAsprrAlaThrleuLeuIleHis 148
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
469 AAGGCTTTCGAA.....TTGCAAAACAGCCGGTCTCTTCCAGA 506
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 GluGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProAs 165
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
507 GCCAAGACAGAGGTGACAACTCGAATACCGCTGCTTCTGTCATCACTCA 556
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 ProThrSerThrAspAsnProGluPheGlnIleValIleSerIleIle 182
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
557 AGCAGTCTTCAACCAAGCAAGAACCACTGCGCACAGCTTGTCCGGC 606
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 YslGluGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 198
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
607 ATGAACGCTGTTCCGAGAGACACACACAGGTGTCACCGCTTACCA 656
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 LeuValGlyValSerGluGluIleThrThrGlyValLysArgLeuTyrGlu 215
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
657 GCTCGAGAGAGGAGGCAAACTCTCTCCAGGCAATCAACGTCAGACAGC 706
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 232
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 CTGTTCAAAAGTCAAGTTTGATTAACATCTACGGCTGCGCACACTCCCTT 756
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 erValThrLysSerLysPheAspAsnLeuTyrGlyArgHisSerLeu 248
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
757 ATCGATGATATCAACCGCTCTCCGATGTCATGATGCGCGGCAAGACAGC 806
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 ProAspGlyLeuMetCArgAlaThrAspValMetIleAlaGlyLysValAl 265
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 TCTGTCATGAGGTACGGGATGCGGATCGGCAAGGCTGCGGTCATACCTCC 856
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 avallilecysglYTrGlyAspValGlyLysGlyCysAlaAlaAlaMet 282
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
857 GTGGCAAGGCGCTGCGCTTATCATCACAGAACTGACCCCAATCTGGCT 906
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 ystThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
907 CTCACAGCTGCCATGGAAGGCTACCAAGTCCGCCGATCGAGAAAGTCGT 956

```

```

||||| ||||| ||||| ::|||
299 LeuInIaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValIva 315
957 CAAGAGATGTCGATATCTTGTACATGACACAGAAACTGCGATATCATCT 1006
||||| ||||| ||||| |||||
315 IserGluIaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
1007 CTGTTGACATGATGGCCGATGAGGATTAAGGCTATTGCGTAACATC 1056
||||| ||| ::||| ||||| |||||
332 etValAspHisMetLargLysMetLysAsnAsnIaIleValCysAsnIle 348
1057 GGCCACTTCGATACGAAATTTGATACAGATGCCCTCATGAAATACCCAGG 1106
||||| ||||| ||||| |||||
349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyProG1 365
1107 CATCAAGCACATCCCAATCAAGCCAGAAATACGACATGTGGGAATCCGAG 1156
::||| ||||| ||||| |||||
365 yValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProG 382
1157 ATGGCCAGCT...ATCTCTCTTGTGCTGAGGCCGCTTCTTAACCTT 1203
:: ||| ||||| ||||| ||||| |||||
382 IuThrLysIaGlyIleIleValLeuIaGluGlyArgLeuMetAsnLeu 398
1204 GGCTGCGCTACAGGTCACCATCTTTGCTTATGTCATGTCACAAA 1253
||||| ||||| ||||| ||||| |||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACACGCGCTGAGCTGACCTCTAGAAAGAGA.....GGAAATC 1297
||||| ||||| ||||| ||||| |||||
415 ngInValIleAlaGlnLeuGlnLeuTrpAsnGluLysAlaSerGlyLysT 432
1298 TCGAGAGAAGATTTCACACTTCGAGACATCTCGATGAGAGAAGTCGCT 1347
||||| ||||| ||||| ||||| |||||
432 yTGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
1348 CGCCTTCACCTCGGATCTCTGATGTCACCTTACAAGCTTACACAG 1395
||||| ||||| ||||| ||||| |||||
449 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLys 464

```









```

381 GAACACATCCGCGCTTCACATGGCCAGATGGTCAAGGCCCCAGCAGG 430
    ||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 pcysthrleuglnalaileuthrtrproaspdlgylgiprosanmet 134
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
431 TTTCGATGATGGTGGTGTATGTACACTCCATCTCCAAAGGCTTCGAA 480
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 leleuaspaspdlgylgiaspalaithrmetleuvalhislysglyval 150
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 TTCGAACAGCCGGTGTCTTCAGAGCCAGACAGAGCTGCACAACTCGA 530
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 tyrglytysalaaglyalavalproaspdpthralsaspserglygl 167
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
531 ATACCGCGCGCTTCCTGCTACACTCAAGCGCTTCACCAAGACAGACA 580
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 uphneatrglylevalleuileuileuargtrpthrleuaglaipropa 184
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
581 ACCATGCGCACACAGTTGCTGCGCGCATGAAGGTGTTCGAGAGACAA 630
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 rglgstprthrleuglnlealaserargileysglyvalthrcluglnthr 200
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
631 ACAACAGGTGTCCACCGCTCTACACAGCTCGAGAGAGGCGCAACTCT 680
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 thrthrlyvalhisargleuylrglumetglnarserglythrleu 217
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
681 CTCCACAGCATCAAGTCACGCGTGTTCACAAAGTCGAGTTCGATA 730
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
217 upheproalaileasnvalasnpservaltthrlyserlyspheaspa 234
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
731 ACATCTACGCGTCCGCGCACCTCCCTTATCGATGATCAACCGTCTCC 780
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 snlystyglycysarghisserleuileaspglyileasnargalethr 250
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
781 GATGTGATGATCGCGCGCAAGACAGCTCTGTCATGGGTACGGCGATGT 830
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 aspyalleuileglyglylyvalalavalalaglytyrglyaspyva 267
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
831 CGGCAAGGGGCGCTCATCCCTCGTGGCGCAAGGCGCGCTTCATCA 880
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
267 lgllysglycysalaglnserleuarglglnglyalargvalylleu 284
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
881 TCACAGACCTGACCCCAATCTGCGCTCTCCAGCTGCCATGGAAGGCTAC 930
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 althrnglileasprroilecysalaleuclnlaalameaspdllytr 300
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
931 CAGTCCGCGCATCGAGGAGTGTCAAGATGTGATATCTTCGTTAC 980
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 glavaltthrthrleuaspdlvalvalglutthralsaspletheth 317
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
981 ATGCACAGAACTGCGATATCATCTCTGTGACATGATGGCCCAAGATGA 1030
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 ralethrgllysnaraspvalillethralsaglnhismetalaargmetl 334
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1031 AGGATTAAGGCTATTGTGCTTACATCGGCGCACTTCGATAAGAAATTGAT 1080
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
334 yshisglnalailevalgllyasnilleglyhispheaspsannglileasp 350
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1081 ACAGATGCGCTCATGAATACCAGGCGATGACACATCCCAATCAAGCC 1130
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 metalaaglyleuvalalaileleproglylleuglylileuglyleuyspr 367
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1131 AGAATACACAACTGGGAATGCCAGATGGCGACGCGATCTCTCTCTG 1180
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 oglnvalhisgluttrpalthrproaspdlgylhisserillelevalleu 384
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1181 CTGAGGCGCGCTTCTTAACCTTGGCTGCGCTACAGGTCAACCATCTTTC 1230
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
384 erglglgllyargleuileuasnleuglyasnalaithrglyhisproserphe 400
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1231 GTTATGTAAATGTCATTACAAACAGACAGCTGCTGAGCTGCACTCTTA 1280
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 valmetserthrserptherhisnglnvalillealaglnilleuglyleuph 417
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1281 CGAAAGAGAGGAATCTCGAGAGAGGTTTACACACTTCGGAAGCATC 1330
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

: ||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 ethrlystpraspdlglnlytrprothrlyvaltytrvalleuprolyshisl 434
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1331 TCGAGAGAAAGTGGCTGGCTCCACCTCGGATCTCTGATGATCCACCT 1380
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
434 euaspdlgylvalalaargleuileuasnspalaleuglyvallyleu 450
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1381 ACAAGCTTTCACACAGAGCGCTGACTACATCAACTCCAGTTCGAGCG 1430
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 thrargleuthrlysglgnalaiglnlytrtyrilleglyvalaspleaspdl 467
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1431 TCCTTACAGTGTGATCGCTTACCGTTAT 1458
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
467 yprotyrlysseraspshistyrargtyr 476
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

seq_name: /cgn2_6/plodata/1/paa/US09_NEW_COMB.pep:US-09-791-537-100163
seq_documentation_block:
; Sequence 100163: Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMIL
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 100163
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor A3
US-09-791-537-100163

alignment_scores:
    Quality: 1645.00      Length: 477
    Ratio: 4.102          Gaps: 1
    Percent Similarity: 84.067      Percent Identity: 65.618

alignment_block:
US-09-759-990-1 x US-09-791-537-100163 ..

Align seg 1/1 to: US-09-791-537-100163 from: 1 to: 485

34 GAGTACAGAAATGCCAGATCAACCTTCAGTTCGGCGGTAAGAACT 83
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9 Asphelelyvalalalaspleuserleuvalalaarhegllyarglysglnl 25
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 TACCTTGTGAGAAAGAAATGCCAGGCTTATGTTGTTGCTGAGCGTT 133
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 ethrleuvalaglnhisglumetproglyleuemetalailearglysglnt 42
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 ATTCCGCTTCTAAGCCATTGAAGGTGTGTCAAGAACTCTGTCTCCAC 183
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
42 yralaglnleaglnproleuvalaglyalargvalthrlyserleu 58
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 ATGACAGTCCAGACAGCGCTTCATCGAGACATCACTCAAGCTTCGTGTC 233
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
59 metthrvalglntthrvalavalleuileuglythrleuvalalaleuglyal 75
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 TGATGTCAGATGGGCTTCTGCAACATCTTCTTACACAAAGATACAGCG 283
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 agluvalargtrpalthrsercysasnlepheserthrclnsphtsalaa 92
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 CTGCTGTATGCTTGTGCGGCCAAGAGCAACAGCAAGAGAGCGCGGT 333
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
92 laalalailealvalaglyproasnnglythrproaspdpthrproglngly 108
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
334 ATCCAGTCTGCGCTGGAAGGCGAAACACTCCCAAGATCTCGGAGAA 383
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

109 ValProValPheAlaTrpIlySglYgluThrLeuGluGluTyrTrpIcy 125  
384 CACATACCGCGCTCTCACATGGCCAGATGTCAA...GGCCCAACAC 427  
125 sThrGlnAlaLeuThrTrpProAsnThrProThrGlyGlyProAsnM 142  
428 AGGTGTGCATATGTTGGTGTGATGCTACACCTCTCATCTCCAAAGGCTTC 477  
142 eTlleuAspAspGlyGlyAspAlaThrLeuLeuValHisIysGlyVal 158  
478 GAATTCCGAACACCGCGCTCTGTCCAGAGCCCAAGACGATGCACACCT 527  
159 GluTyrGluIlyAspIlyIysValProSerValAspThrAlaGluSer 175  
528 CGAATACCGCGCTGTTCTTGTCACTACAGACGAGTCTTCAACCAAGACA 577  
175 pGluHisArgAlaIleLeuGluIleuLeuThrArgThrValGlyGluSer 192  
578 ACACACCATGCGCACACAGTTGCTGCCGGCATAGACGGTCTTCCCAAG 627  
192 roGluIysTrpThrGluLeuAlaSerGluIleArgGlyValThrGluGlu 208  
628 ACACACCAAGGCTGCCACCGCGCTTCACAGCTCGAAGACCAAGGCGCAACT 677  
209 ThrThrThrGlyValHisArgLeuIyrGluMetHisArgAspIlyThrIle 225  
678 CCTCTTCCACGCCATCAACAGTCACAGACGCTGTACAAAGTCCAAAGTTG 727  
225 IueupPheProAlaIleAsnValAsnAspAlaValThrIysSerIysPhe 242  
728 ATPACATCTACGGCTGCGCGCCACTCCCTTATGATGGATTCACACGCTGT 777  
242 sPAsnIysTyrGlyCysArgHisSerLeuIleAspIlyIleAsnArgAla 258  
778 TCCGATGATGATGTCGGCGGCAAGACAGCTCGATGAGGTATACGCGCA 827  
259 ThrAspValLeuIleIylGlyIlyThrAlaValCysGlyTyrGlyIys 275  
828 TGTGCGCAAGGCGTGGCTCAATCCCTCCGTGGCCAAAGGCGCTCGGCTTA 877  
275 pAlaGlyIysGlyCysAlaGluSerLeuArgGlyGlnGlyAlaArgValI 292  
878 TATATCAAGAACTCCACCCAACTCGGCTCTCCAGGCTCCCATGGAAGC 927  
292 IeIleThrGluIleAspProIleCysAlaLeuGluAlaIleMetAspGly 308  
928 TACACAGTCCGCGCATCGAGAAAGTCGCAAGATGGATATCATCTGCT 977  
309 pHeGluValThrThrLeuAspGluValValAspIysAlaAspIlePheVal 325  
978 TACATGCAAGAAACTGCGATATCATCTCTGTGCATGATGGCCCA 1027  
325 ThrThrThrGlyAsnIlyAspIleIleMetAlaIysAspMetAlaIysM 342  
1028 TGAAGATTAAGGCTATTTGTCGGTACACATGGCCACTTCATACGAAAT 1077  
342 eTlysHisGlnAlaIleValGlyAsnIleGlyHisPheAspAsnGluIle 358  
1078 GATACAGATGGGCTATGAAATACCAAGGATCAAGACATCCCAATCAA 1127  
359 AspMetAlaGlyIleuAlaGluThrProGlyIleValIysAspGluValIly 375  
1128 GCGAGATACGACATGTGGGAATTCACAGATGGCCACGCTATCTCTTC 1177  
375 sProGluValHisThrTrpThrTyrProAspGlyIysValLeuIleVal 392  
1178 TTGCTAGAGGCGCGCTTTAACTTTGGCTGGCGCTACAGGTCAACCATCT 1227  
392 euserGlnIlyArgLeuAsnLeuGlyAsnAlaThrGlyHisProSer 408  
1228 TTTCGTATGTCATGTCATTCACAAACACAGACACTGGCTACAGTCACT 1277  
409 PheValMetSerAsnSerPheAlaAspGluThrIleuAlaGluIleIleu 425

[illegible]



```

120 ValProValPheAlaTrpLysGluThrLeuGluGluTrpTrpAl 136
136 CACATACGCGCTCCACATGGCCAGATGGTCAAGGCCA...CAGCAGG 430
136 AlaAlaGluMetLeuThrTrpProAspProAspLysProAlaAsnMetI 153
431 TTGTCGATGATGGTGGTGTGTACACTCCCTCATCTCCAAAGGCTTCGAA 480
153 IeLeuAspAspLysGlyAspAlaThrMetLeuValLeuArgGlyMetGln 169
481 TTCGAAACAGCGGCTGCTTCACAGGCCAACAGAACTGACAACTCGA 530
170 TyrGluLysAlaGlyValAlaProProAlaGluGluAspAspProAlaG1 186
531 ATACCGCTGCGCTTCTGTACACTCAAGAGCTCTTCAACCAAGACAGA 580
186 uTrpLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLysA 203
581 ACCACTGGCACACAGTGTGCGCGCATGAAAGGCTGTTCGAGAGACA 630
631 ACAAGAGGTGTCACCGCTCTACCGAGCTCGAGAGAGGCGCAACTCCT 680
220 ThrThrGlyValLeuArgLeuTrpGlnPheAlaAlaAlaLysAspLeuAl 236
681 CTTCGCCACCATCAACGTCACGAGCGCTGTTCACAAAGTCCCAAGTTCGATA 730
236 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 253
731 ACATCTACGGCTGCCGCCACCTCCCTTATCGATGATCAACCGTGTCC 780
253 snLysTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 269
781 GATGTCATGATGCGGCGCAAGACAGCTCTGTCATGGTTCACGCCATGT 830
270 AspAlaLeuIleGlyGlyLysLysValLeuIleCysGlyTyrGlyAspVa 286
831 CGGCAAGGGCTGCGCTCAATCCCTCCGTCGCGCAAGCGCTCGGCTATCA 880
286 IGlYysGlyCysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 303
881 TCACAGAACTCGACCCCAATCTCGCTCTCCAGGCTGCCATGAGAGCTTAC 930
303 aThrGluIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 319
931 CAGTCCCCCGATCGAGAACTCTCGCAAGAGATGATGATCTTCCTTAC 980
320 AspValValThrValGluGluAlaIleGlyAspAlaAspIleValAlaThr 336
981 ATGCACAGAACTCGCATATCATCTCTGTGATCATGATGAGCCCATGA 1030
336 GlAlaThrLysAsnLysAspIleIleMetLeuGluIleLysAlaMetL 353
1031 AGATTAAGCTATGTGCGTAACTCGGCACTTCGATTAAGAAATTTGAT 1080
353 yAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnIleAsp 369
1081 ACAGATGGCTCATGAATACCCAGGATCAAGACATCCCAATCAAGCC 1130
370 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysPr 385
1131 AGAATACACATGTGGGAATTCACAGAT...GGCCACGCTATCCCTCTC 1177
385 ogLValAspLeuTrpThrPheGlyAspThrGlyArgSerIleIleValL 402
1178 TTGCTGAGGCGCGCTTAACTGTGCGCTACAGGTCAAGCTCACTCT 1227
402 euSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 418
1228 TTCGTTATGTCATGTCATCAACAACGACATGCTGCTCAGCTCAGCT 1277

```

```

419 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 435
1278 CTACGAAAGAGAGGAATATCGAGAGAAAGCTTTACACACTTCGGAAGC 1327
435 uTrpThrLysAsnAspLysLysValLysAspAsnGluValLysArgLeuProLysH 452
1328 ATCTGATGAAGAAGTGCCTCGCTCCACCTCGGATCTCTCGATGTCAC 1377
452 ILeuAspGluLysValAlaArgIleHisValGluAlaLeuGlyLysHis 468
1378 CTTCACAAAGCTTACACAGAGAGCGTGCATACATCAACGTTCCAGTTGA 1427
469 LeuThrLysLeuThrLysGluGlnAlaGluTrpLeuGlyValAspValG1 485
1428 GGGCTCTTACAGTCTGATGCTTACCGTTAT 1458
485 uGlyProTyrLysProAspHisTyrArgTyr 495

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-791-537-37085
seq_documentation_block:
: Sequence 37085, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: NUMBER OF SEQ ID NOS: 2001-02-22
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 37085
: LENGTH: 485
: TYPE: PRT
: ORGANISM: Petroselinum crispum
: US-09-791-537-37085

alignment_scores:
Quality: 1544.00 Length: 488
Ratio: 3.879 Gaps: 6
Percent Similarity: 81.557 Percent Identity: 63.115

alignment_block:
US-09-759-990-1 x US-09-791-537-37085 ..

Align seg 1/1 to: US-09-791-537-37085 from: 1 to: 485

19 ACTGGTCTCCATTCGAGTACAGAAATGGCCAGATCAACCTCATGTCT 68
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
8 ThrAlaAlaGlyArgGluTrpLysValLysAspMetSerLeuAlaAspPh 24
69 CGGCGCTAAGGAATTACCTTCTGAGAAAGAAAGAAATGCCAGTCTTATGG 118
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
24 eGlyArgLeuGluLeuGluValLeuAlaGluValAlaGluMetProGlyLeuMetS 41
119 TTCTTCGTAGAGCTTATCCGCTTCTTAACCATTTGAAGGGTGTGAGATC 168
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41 eCysArgThrGlnPheGlyProSerGlnProPheLysGlyAlaArgIle 57
169 TCTGGTTCCTCCATGACATGACATGACATGACATGACATGACATGACAT 218
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
58 ThrGlySerLeuHisMetThrIleGlnThrGlyValLeuIleGluThrIle 74
219 CACAGCTTGTGCTGATGTCAGATGGCTTCTCGACAACTTCTCTTA 268
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91
269 CACAGATACAGCGCGCTGCTGCTATGCTGTTCGGCCCAACAGGACACA 318
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 hGlnAspHisAlaAlaIleAlaIleAla.....'..... 100

```



```

|||||
74 uThrlAlaLeuGlyAlaGluValArgTyrPcysSerCysAsnIlePheSer 91
269 CACAAGATATACAGCCGCTGCTGATCGTTGTGGCCACACAGGACACCA 318
91 hrgInaPheHisAlaAlaAlaIleAlaIleAlaArgAspSerAlaIa..... 105
319 GAGAACGACGCGGTATCCCACTCTTCGCTGGAAGGGGAAACACTGCC 368
106 .....ValPheAlaIleArgPheGlyGluThrLeuGlu 115
369 AGAATACGTGGGAGAACATACCGGCTCTCCATCGCCAGATGCTGCAAG 418
115 nglutYrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGly 132
419 GCCCAGACAGAGTTGTGATGATGTTGATGATACACTCTCTCATCTCC 468
132 LyrProAspLeuIleValAlaAspArgGlyAspAlaThrLeuLeuIleHis 148
469 AAGGCTTCGAA.....TTCGAAACAGCGCGGTGCTTCAGCA 506
149 GluGlyValIysAlaGluGluIlePheGluIysThrGluValProAs 165
507 GCCAAGCAGAGCTGACAACTCGATACCGCGCTTCCTGCTACACTCA 556
165 pProThrSerThrAspAsnProGluPheGluIleValLeuSerIleIle 182
557 AGCAGCTCTTCACCAACCAAGACAAACACACACTGACACAGTTGCTGCGCG 606
182 YsgIuGlyLeuGluIleValAspProIysIysThrHisLysMetLysGluArg 198
607 ATCAACGGGTGTTCCGAGAGACAAACAGAGTGCACCGCTCTACCA 656
199 LeuValGlyValSerGluGluThrThrGlyValIysArgLeuIyrGlu 215
657 GCTCGAGACAGAGGCAAACTCTCTCCAGCCATCAAGCTCAACGACGAG 706
215 nMetGluGlnIAsnGlyThrLeuLeuPheProAlaIleAsnValAspS 232
707 CTGTACAAAGTCCAAAGTTCGATACATCTAGCGCTGCGCCACTCCCT 756
232 eValAlaThrLysSerLysPheAspAsnLeuYrGlyCysArgHisSerLeu 248
757 ATCGATGATATCAACCTGCTCCGATGTCATGATCGCGCGGCAACAGC 806
249 ProAspIleuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
807 TCTCGATAGGTTACCGCATGTCGCGAAGGCTCGCTCAATCCCTCC 856
265 aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaIleMet 282
857 GTGGCCAAAGCGCTCGCTTATCATACAGAACTGACCCCAATCTGCGCT 906
282 YsThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
907 CTCCAGGCTCCCATGGAAGGCTACAGGTCGCCGCGCATCGAGGAAGTGT 956
299 LeuGlnAlaLeuMetGluGlyLeuGluValLeuThrLeuGluAspValVa 315
957 CAAGATGTCGATATCTCTTACATGCACAGAACTGCGTATTCATCT 1006
315 lSerGluAlaAspIlePheValThrThrThrGlyAsnLysAspIleIle 332
332 eValAlaPheIleMetArgLysMetLysAsnAlaIleValCysAsnIle 348
1057 GCGCCTTCGATTAACGAATTGATACAGANTGCTCATGAATACCCAG 1106
349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTrpGlu 365
1107 CATCAAGCATCCCAATCAAGCCAGAAATAGACATGTGGGAATCCAG 1156
|||||

```

```

365 yValLysArgIleThrIleLysProGluThrAspArgTrpValPheProG 382
1157 ATGGCCAGCCT..ATCCCTCTCTTGCATGAGGCGGCTCTTACACTT 1203
382 lUThrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
1204 GCGCTCGCTACAGGTCAACCATCTTTCGTATGTCATTCATTCACAAA 1253
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACTCGCTCAGCTCGACTCTTACGAAAGAGA.....GGAATC 1297
415 nglValIleAlaGluIleGluLeuIleuTrpAsnGluLysAlaSerGlyLys 432
1298 TCGAGAGAGAGGTTTACACACTCCGAGAGCATCGATGAGAGAGTGC 1347
432 YrGluLysAlaValIyrValIleuProLysHisLeuAspGluLysValAla 448
1348 CGCCTCAACCTCGATCTCTCGATGTCACACTTACAAAGCTTACACAGAA 1397
449 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAs 465
1398 GAGGCTACTCATCAACAGTTCAGATTGAGGCTCTTACAACTGATG 1447
465 pLisSerAspTyrValSerIleProIleGluGlyProTyrLysProPro 482
1448 CTACCGCTTAT 1458
482 lTyrArgTyr 485

seq_name: /cgn2_6/prodata/1/pae/us09_NEW_COMB.pep:us-09-935-625-2153
seq_documentation_block:
; Sequence 2153, Apolipoprotein US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF C
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 2153
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..485
; OTHER INFORMATION: Ceres Seq. ID no. 3439072
US-09-935-625-2153

alignment_scores:
Quality: 1536.00 Length: 487
Ratio: 3.898 Gaps: 4
Percent Similarity: 80.903 Percent Identity: 62.012

alignment_block:
US-09-759-990-1 x US-09-935-625-2153 ..
Align seg 1/1 to: US-09-935-625-2153 from: 1 to: 485

19 ACTGTCCTCATTCGATACAGAAATTCGCGACATCACTCATCTTCT 68
|||||
8 ThrSerSerGlyArgGluTyrIleLysValIysAspMetSerGlnAlaAsp 24
69 CGCGCTTAAGGAACTTACCTTTCGTGAGAAGGAAATGCCAGGTCTTATG 118
|||||
24 eGlyArgLeuGluIleuGluIleuAlaGluValGluMetProGlyLeuMet 41
119 TTTCTTCGAGCGTTATTCCTTCAAGCCATTGAAGGTGTGCAAAATC 168
:: |||

```

```

41  IacysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57
169  TCCTGCTCCCTCCACATGACAGTCACAGACCGCTCATCGACACACT 218
    ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58  ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGlnThr 74
219  CACAGCTCTGGTGGATGTCAGATGCGCTCTCTGCAACATCTTCTCTA 268
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
74  uThrAlaLeuGlyAlaGluValArgTyrCysSerCysAsnIlePheSer 91
269  CACAGATACAGCCGCTGCTGATGCTGTCGGCCACAGCAGCACCA 318
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
91  hrGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319  GAGAAAGCAGCCGGTATCCCATCTTCTGCTGGAAGGCGCAACACTCC 368
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
106  .....ValPheAlaIleTyrLysGlyIleThrLeuGln 115
369  AGAATACGTGGAGAACACATACCGGCTCTACATGGCCAGATGGTCAAG 418
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
115  nGluTyrTyrTyrCysThrGluArgAlaLeuAspTyrPheLysGly 132
419  GCCCAGCAGAGGTGTGATGATGATGATGATGATGATGATGATGATG 468
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
132  LyrProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuHis 148
469  AAGGCTTCGAA.....TTCGAAACAGCCGGTGTGTCAGA 506
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
149  GluGlyValIleGluGluGluIlePheGlnIleThrGlyIleValPro 165
507  GGCACAGACAGCTGACACCTGATACCGCTGCTGCTGCTGCTGCTGCT 556
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165  pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLeu 182
557  AGCAGCTCTCAACCAAGAACAGACACAGCAGCAGCAGCAGCAGCAG 606
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
182  ySGLuGlyLeuGlnIleValAspProLysLysTyrHisLysMetLysGlu 198
607  ATGAACGGTGTTCGAGAGACACACAGCAGCAGCAGCAGCAGCAGCAG 656
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
199  LeuValGlyValSerGluGluIleThrThrGlyValLysArgLeuTyrG 215
657  GCTGAGAGAGAGGCGCAACTCTCTCCAGCAGCAGCAGCAGCAGCAG 706
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
215  nMetIleGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 232
707  CTGTACAAAGTCAAGTTCGATACATCTACAGGCTGCGCCACTCCCT 756
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232  ervaLThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
757  ATCGATGGTATCAACCGCTGCTCGATGTCATGATGCGGCGCAAGAC 806
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
249  ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVal 265
807  TCTGCTCAGTGGTACGGGATGCGGAGTGGCAGGCTGCGCTCAATCC 856
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
265  aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetL 282
857  GTGGCAGAGGCGCTGCTGATATCATCAGAACTGCAACCATCTGCGCT 906
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
282  ySThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
907  CTCAGAGCTGCATGAGAGGCTACAGAGTCCCGCATCGAAGAGTGGT 956
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
299  LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGlnAspVal 315
957  CAAGGATGTCGATATCTGCTTACATCAGACAGAACTGCCATATCACT 1006
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
315  lSerIleAlaAspIlePheValThrThrThrGlyAsnLysAspIleLeu 332
1007  CTGTGACATGATGGCCAGCAGATGAGAGATGATGATGATGATGATG 1056
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
332  etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348

```

```

1057  GCCCATCTTGATACGAAATGATACAGATGCGCTCATGAAATCCACAG 1106
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
349  GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGlnThrTyrPro 365
1107  CATCAAGCAGACATCCCAATCAAGCAGAAATACAGATGCGGATCCGAG 1156
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
365  yValLysArgIleThrLeuLysProGlnThrAspArgTyrValPhePro 382
1157  ATGGCCACGCT...ATCTCTCTCTCTGCTGAGGCGCGCTCTTAACCT 1203
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
382  lThrLysAlaGlyIleIleIleValLeuAlaGluGlyArgLeuMetAsn 398
1204  GCTGCGCTTACAGCTGACACCATCTTCTGTTATGTCATGATTCACAAA 1253
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
399  GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThr 415
1254  CGAGACATCGTCGATCGACCTTCGACAAAGAGAGAGAGAGAGAGAG 1297
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
415  nGlnValIleAlaGlnLeuGluLeuTyrPheAsnGluLysAlaSerGly 432
1298  TCGAGAGAGAGGTTCACACTTCCGAGACATCTCGATGAGAGAGTGGCT 1347
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
432  yGluLysLysValIleValLeuProLysHisLeuAspGluLysValAla 448
1348  GCGCTCCACCTCGGATCTCGATGTCACACCTTACAAAGCTTACACAGA 1397
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
449  LeuLeuHisLeuGluLysLeuGlyAlaArgLeuThrLysLeuSerLys 465
1398  GCAGCGTGAATACATCAAGCTTCCAGTTGAGAGGCTTACAAAGTGTGAG 1447
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
465  pGln**AspTyrValSerIleProIleGluGlyProTyrLysProPro 482
1448  CTTACCGTTAT 1458
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
482  lSTyrArgTyr 485

```

```

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-935-625-2152
seq_documentation_block:
; Sequence 2152, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 2152
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..507
; OTHER INFORMATION: Ceres Seq. ID no. 3439071
US-09-935-625-2152

```

```

alignment_scores:
    Quality: 1536.00      Length: 487
    Ratio: 3.898          Gaps: 4
    Percent Similarity: 80.903      Percent Identity: 62.012

```

alignment\_block:

US-09-759-990-1 x US-09-935-625-2152 ..

Align seg 1/1 to: US-09-935-625-2152 from: 1 to: 507

```

19  ACTGGGCTCCATTCGATGACAGAAATGCGAGATCAACCTCCAGTGTCT 68
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
30  ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspH 46

```

```

69 CGGCCGTAAGAACTTACCTGTCGAGAAGAAATGCCAGTCTTATGG 118
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
46 eGlyArgLeuGluLeuGluLeuAlaGluValGluMetProGlyLeuMet 63
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
119 TTCTTGAGTACGCTTATTCGCTTCTAAGCCATTGAAGGTCAGATC 168
   :||| :||| :||| :||| :||| :||| :||| :||| :|||
63 LacysArgThrGluPheGlyProSerGlnProPheGlyAlaArgIle 79
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
169 TCTGGTCCCTCCACATGACATGACAGACGCGCTCTCATCGAGACT 218
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
80 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThr 96
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
219 CACAGCTCTTGTCGATGTCAGATGGCTTCTCCATCATCTTCTCTA 268
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
96 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 113
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
269 CACAAGATACGCGCTGCTGCTATGCTGTGCGCCCAACAGCACCA 318
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
113 hrgInaSPHisAlaAlaIleAlaIleAlaArgAspSerAlaAla.... 127
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
319 GAGAAGCAGCGCGTATCCCACTCTTCGCTGGAAGGCGCAACACTCC 368
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
128 .....ValPheAlaTrpIlyGlyIleThrLeuGlu 137
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
369 AGAATAGTGGAGAACACATACCGCGCTCTACATGGCCAGATGTCAG 418
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
137 ngIuTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyIlyG 154
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
419 GCCCAGCAGAGGTGTCATGATGTCGTCGATGTCATCAGCTCATCTCC 468
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
154 LyProAspLeuIleValAspArgIlyGlyAspAlaThrLeuLeuHis 170
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
469 AAGGCTTCGAA.....TTCGAAACAGCGGCTGTCGTCGAGA 506
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
171 GluGlyAllyAlaGluGluLeuPheGluIlyThrGlyGlnValProAs 187
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
507 GCCACAGAAAGCTGACAACTCGAATACCGCTGCTTGTGTCACACTCA 556
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
187 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLe 204
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
557 AGCAGGCTTCAACCAAGACAAAGAACACTGGCACACAGTGTGCGCGC 606
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
204 ysgIuGlyLeuGlnValAspProIlyLysTrpHisIlysmetLysGluArg 220
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
607 ATGAACGCTGTTCCGAGAGACAAACAGAGTGTCCACCGCTCTACCA 656
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
221 LeuValGlyAlaSerGluGluThrThrGlyValLysArgLeuTrpGlu 237
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
657 GCTCGAAGAGAGGCAACTCTCTTCCAGCATCAACGTCACAGAG 706
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
237 nmetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 254
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
707 CTGTCAAGTCCAGTTCGATACATCTAGCGCTGCCGCCACTCCCTT 756
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
254 eValIleThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 270
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
757 ATCGATGATCAACCGTGTCCGATGTCATGATCGCGGCAAGACAGC 806
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
271 ProAspIlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 287
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
807 TCTCGTATGGGTTACGGCGATGTGGCAAGGCTGCGCTCAATCCCTCC 856
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
287 aValIleCysGlyTrpGlyAspValGlyLysGlyCysAlaAlaIleMetL 304
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
857 GTGGCCAAAGGCGCTCGCTATCATCACAGAACTGCAACCAATGCGCT 906
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
304 ysthAlaGlyAlaArgValIleValIleGluIleAspProIlyCysAla 320
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
907 CTCACGAGCTGCATGAGAGGCTACAGAGTCCCGCGCATGAGAGAGTGT 956
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
321 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGlnAspValVa 337

```

```

957 CAAGGATGTCATATCTTGTTCATGTCACAGGAATGCCAGATTCATCT 1006
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
337 lserGluAlaAspIlePheValThrThrGlyAsnLysAspIleLeu 354
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1007 CTGTTCATATGATGCGCCCAATGAAGGATAGGCTTATGTCGTAATC 1056
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
354 eValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 370
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1057 GGCCATTCGATTAACGAATGATACAGATGGCGCTCATGAATACCCAG 1106
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
371 GlyHisPheAspAsnGlnIleAspMetLeuGlyLeuGluThrTrpProG 387
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1107 CATCAAGCAATCCCAATCAAGCCAGAAATACGATGTGGAAATCCAG 1156
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
387 yValIlyArgIleThrIleLysProGlnThrAspArgTrpValPheProG 404
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1157 ATGGCCAGGCT...ATCCCTCTTGTGTCGAGGCGCGCTTCTTAACCTT 1203
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
404 luthrLysAlaGlyIleIleValIleValAlaGluIlyArgLeuMetAsnLeu 420
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1204 GGCCTGCTACAGGTCACCCATCTTTCGTTATGTCATGTCATTCACAAA 1253
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
421 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 437
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1254 CCAGACACTGCTCAGCTGACCTTACGAAAAGAA.....GGAATC 1297
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
437 ngInValIleAlaGlnLeuGluLeuThrAspAsnGlyLysAlaSerGlyLys 454
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1298 TCGAGAGAGAGTTCATCACTTCGAAAGCATCTGATGAGAAGTGGCT 1347
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
454 yrgIuLysLysValTrpValLeuProLysHisLeuAspGlyLysValAla 470
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1348 CGCCTCCACCTCGGATCTCTGATGTCCACTTCAAAAGCTTACACAAA 1397
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
471 leuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAs 487
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1398 GCAGGCTGACTACATCAACGTTCCAGTTGAGGCTCTTACAGTGTGATG 1447
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
487 pGln**AspTrpValSerIleProIleGluGlyProGlyLysProProH 504
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
1448 CTTCACGCTTAT 1458
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||
504 lstyArgTrpTyr 507
   :||||| :||||| :||||| :||||| :||||| :||||| :|||||

```

```

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-791-537-49252
seq_documentation_block:
: Sequence 49252, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomedix, Inc.
: APPLICANT: Debe, Derek
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791.537
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 49252
: LENGTH: 485
: TYPE: PRT
: ORGANISM: Nicotiana sylvestris
US-09-791-537-49252

alignment_scores:
  quality: 1533.00      length: 482
  ratio: 3.951          gaps: 4
  Percent Similarity: 80.498      Percent Identity: 63.278

alignment_block:
US-09-759-990-1 x US-09-791-537-49252

```



Align seg 1/1 to: US-09-791-537-49252 from: 1 to: 485

```

34 GAGTACAGAAATGGCGACATCAACCTCCATGTTCTGCGCGGTAAAGAACT 83
   |||||
13 GluTrpLysValIleAspMetSerGlnAlaAspPheGlyArgLeuGlnI 29
   : |||||
84 TACCTTGTGGAAGAAATGCCAGCTCTTATGGTCTTCTGTCGAGCTT 133
   : |||||
29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgTrpGlu 46
   : |||||
134 ATTCGCTTCAAGCCATGAAAGGTGTGTCAAAATCTGTGTTCCCTCCAC 183
   : |||||
46 heGlyProSerGlnProPheLysGlyAlaLysIleThrGlySerLeuHis 62
   : |||||
184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACACTCACAGCTTTGGTGC 233
   : |||||
63 MetThrIleGlnThrAlaValIleGlnThrLeuThrAlaLeuGlyAla 79
   : |||||
234 TGATGTGATGGGCTTCTGCAACATCTCTTACACAGATACAGCGG 283
   : |||||
79 aGluValArgTrpCysSerCysAsnIlePheSerThrGlnAspHisAla 96
   : |||||
284 CTGCTGCTATCGTTGTGCGCCCAACAGCCACAGAGAGAGAGAGAGAG 333
   : |||||
96 IaAlaIleAlaIleAlaArgAspSerAlaAla..... 105
   : |||||
334 ATCCAGTCTGCTGCGTGAAGAGGCGAAAGACACTCCCAAAATCTGGGAGA 383
   : |||||
106 .....ValPheAlaTrpLysGlyIleThrLeuGlnIleGlyTrpTrpCys 120
   : |||||
384 CACATACAGCGGCTCTCAGATGGCGGAGATGGTCAAGCCACAGAGGTG 433
   : |||||
120 sThrGlnArgAlaLeuAspTrpGlyArgIleProGlyArgProAspLeuIle 137
   : |||||
434 TCGATGATGGTGTGATGTACACTCTCTCATCTCCAGAGGCTTC..... 477
   : |||||
137 aAspArgSerGlyIleAspAlaThrLeuLeuIleHisGlyValIleLysAla 153
   : |||||
478 .....GAATTCGAAACAGCGCGGTGCTGACAGAGAGAGAGAGAGAG 521
   : |||||
154 GluGlnGlnPheAlaLysAsnGlyThrIleProAspProAsnSerThrAs 170
   : |||||
522 CAACCTCGAATACCGGCTGCTTGTGTACACTCAAGCAGAGTCTTCAACC 571
   : |||||
170 pAsnAlaGlnPheGlnLeuValIleThrIleLeuLysGlnSerLeuIle 187
   : |||||
572 AAGACAAGAACCACTGGCACACAGTGTGCTGCGGATGAAAGGTGTTTC 621
   : |||||
187 hAspProLeuLysTrpThrLysMetLysGluArgLeuValGlyValSer 203
   : |||||
622 GAAAGACAAACAGAGTGCACCGGCTTACACAGCTGAGAGAGAGG 671
   : |||||
204 GluGlnThrThrThrGlyValLysArgLeuTrpGlnMetGlnAlaAsnGln 220
   : |||||
672 CAACACTCTCTTCCAGCATCAACAGTCAAGCAGCGTGTACAAAGTCA 721
   : |||||
220 yThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSer 237
   : |||||
722 AGTTGATTAACATCTACGGCTGCGCGCACTCCCTTATCATGATGATCAAC 771
   : |||||
237 ySphAspAsnLeuTrpGlyCysArgHisSerLeuProAspGlyLeuMet 253
   : |||||
772 CGGTGTCGATTCATGATGAGGCGGAGAGAGAGCTGCTGTCATGGGTTA 821
   : |||||
254 ArgAlaIleThrAspValMetIleAlaGlyLysValAlaLeuValAlaGly 270
   : |||||
822 CGGCGATGTGGGCAAGGGCTGCGTCAATCCTCGTGGCCAAAGGCGCTC 871
   : |||||
270 rGlyAspValGlnLysGlyCysAlaAlaLeuLysGlnAlaGlyAla 287
   : |||||
872 GCGTATCATACAGAACTGCAACCCAAATGCTGGCTCTCAGAGGTGCATG 921
   : |||||

```

```

287 rValIleValThrGluIleAspProIleCysAlaLeuGlnAlaThrMet 303
922 GAAGCTACAGAGTCCGCCCATCGAGAGAGTGTCAAGATGTGCATAT 971
   |||||
304 GluGlyLeuGlnValLeuThrLeuGlnAspValValSerAspValAsp 320
   : |||||
972 CTTCGTTACATGACAGAAACTGCGATTCATCTCTGTTGACATGATGG 1021
   : |||||
320 ePheValThrThrThrGlyAsnLysAspIleIleMetValAspHisMet 337
   : |||||
1022 CCCAGATGAAGGATTAAGGCTATGTGCGGTAAACATCCGCCACTTCAT 1071
   : |||||
337 rGlyMetLysAsnAlaIleValLysAsnIleGlyHisPheAspAsn 353
   : |||||
1072 GAAATGTATACAGATGAGCTCATGAAATACCCAGCATCAACACATCCC 1121
   : |||||
354 GluIleAspMetLeuGlyLeuGlnThrTrpProGlyValLysArgIle 370
   : |||||
1122 AATCAAGCCAGAAATACAGATGTGGAAATTCACAGATGCCAGCT...A 1168
   : |||||
370 rIleLysProGlnThrAspArgTrpValPheProAspThrAsnSerGly 387
   : |||||
1169 TCCCTCTTCTTGTGAGAGCGCGCTTCTTACCTTGGCTGGCTACAGGT 1218
   : |||||
387 lIleValLeuAlaGlnGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
   : |||||
1219 CACCCATCTTCTGTTATGTATCATATCATTCACAAACAGACACTGCTCA 1268
   : |||||
404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAla 420
   : |||||
1269 GCTGACCTGTACGAAAGA...GGAATCTCGAAGAAAGATT 1312
   : |||||
420 nLeuGlnLeuTrpAsnGlnLysSerSerGlyLysTrpGlnLysVal 437
   : |||||
1313 ACACACTTCGAGCATCTGATGAGAAGAGTCTGCTGCTCACCTCAGCGA 1362
   : |||||
437 yThrLeuProLysHisLeuAspGlnLysValAlaAlaLeuHisLeuGly 453
   : |||||
1363 TCTCTGATGTCCACTTACAAAGCTTACACAGAGAGAGCTGACATACAT 1412
   : |||||
454 LysLeuGlnLysLeuThrLysLeuSerLysAspGlnAlaAspTrp 470
   : |||||
1413 CAACGTTCCAGTTGAGGCTCTTACAACTGATGTTACCTTAT 1458
   : |||||
470 eSerValProValGlnGlyProLysProAlaHisTrpArgTrp 485
   : |||||

```

seq\_name: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:US-09-791-537-37086

seq\_documentation\_block:

```

; Sequence 37086, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILIES
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37086
; LENGTH: 485
; TYPE: PRF
; ORGANISM: Trilicium aestivum
US-09-791-537-37086

```

alignment\_scores:

```

Quality: 1524.00 Length: 487
Ratio: 3.868 Gaps: 4
Percent Similarity: 80.903 Percent Identity: 61.807

```

alignment block:  
US-09-759-990-1 x US-09-791-537-37086 ..

Align seg 1/1 to: US-09-791-537-37086 from: 1 to: 485

```

19 ACTGGTCATTCAGTACAGAAATTCGCGAGTCAACCTCCATCTCTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluTyrTyrValLysAspLeuPheGlnAlaAsp 24
69 GCGCCGTAAGAACTTACCTTGTGTAAGAAATTCGAGTCTTATGG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
24 eGlyArgGluLeuGluLeuAlaGluValGluMetProGlyLeuMet 41
119 TTCTTGAGAGGGTTATCCGCTTCTAAGCATTTGAAGGTTGTCAGATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
41 LAspArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57
169 TCTGGTCCCTCCATGACATGACATGACATGACATGACATGACATG 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThr 74
219 CACAGCTCTTGGTGTGATGATGATGATGATGATGATGATGATGATG 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
74 ThrAlaLeuGlyAlaGluValArgTyrPheLysSerCysAsnIlePhe 91
269 CACAAGATACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
91 eGlnHisPheHisAlaAlaAlaAlaAlaAlaArgAspSerAlaAla 105
319 GAGAAGCCAGCCGCTGATCCAGTCTGCTGCTGCTGCTGCTGCTGCT 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 .....ValPheAlaThrPheGlyGluThrLeuGlu 115
369 AGATATCTGGAGAAACATATCCGCTGCTGCTGCTGCTGCTGCTGCT 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
115 uGluTyrTrpPheCysThrGluArgCysLeuAspTrpGlyValGly 132
419 GCCCAGCAGCAGTGTGATGATGATGATGATGATGATGATGATGATG 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
132 LysProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuHis 148
469 AAGGGCTTC.....GAATTCGAAGACGCGGTGCTGTGTCAGAA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
149 GluGlyValLysAlaGluGluGluPheGluLysSerGlyLysValPro 165
507 GCCAAGCAAGACTGACATCCCTGATACCGCTGCTGCTGCTGCTGCT 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
165 PProGluSerThrAspAsnProGluPheLysIleValLeuThrIle 182
557 AGCAGTCTTCCAAACCAAGCAAGCAAGCAAGCAAGTGTGCTGCGCG 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
182 rGAspGlyLeuLysThrAspAlaSerLysTyrArgLysMetLysGluArg 198
607 ATGAAACGGTGTTCGGAAGAGACAAACAGAGTGTCCACCGCTTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
199 LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTyrG 215
657 GCTCGAGAGAGGAGGCAACCTCTTCCACGACATCAACGTCACAGAG 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
215 nMetGlnGluSerGlyThrLeuPheProAlaIleAsnValAsnAsp 232
707 CTGTACAAAGTCCAAAGTTCATTAACATTCAGGCTGCGCGACACCT 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232 eValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLe 248
757 ATCCATGATGATCAACCGTCTTCGAGTGTATGATGCGCGCAAGACAG 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyValAl 265
807 TCTGTCATGAGGTATGCGCATGCGGAGAGGCTGCGCTCATCTCCCTCC 856
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
265 aValValCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaAlaLe 282

```

```

857 GTGGCCAAAGGCGCTGCGCTTATCATCAGACAGACTGACCAATTCGCGCT 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
282 YsgInAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
907 CTCGAGGCTGCCATGAAAGGCTTACAGGCTCCGCGCATGAGAAAGTCGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
299 LeuGlnAlaLeuMetGluGlyIleGlnIleLeuThrLeuGluAspValVa 315
957 CAAGATGATGATCTTCTGATCATGACAGCAAGAAATCGCATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
315 LSerGlnAlaAspIlePheValThrThrGlyLysAsnLysIleIleMet 332
1007 CTGTGATCATGATGCGCCAGATGAAAGTAAAGGATATTTGCGGTAATC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
332 eValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
1057 GCGCATCTGATTAACGAAATTTGATACAGATGCGCTCATGAAATCCAGG 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
349 GlyHisPheAspAsnGluIleAspMetAsnGlyLeuGluThrTyrPro 365
1107 CATCAAGCAGATCCCAATCAAGCCAAATACAGATGCGGAAATCCAG 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
365 YValLysArgIleThrIleLysProGlnThrAspArgTyrPheProG 382
1157 ATGGCCAC...GCTATCCTCCTTCTTGTGAGGCGCGCTTCTTAACCTT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
382 LThrLysThrGlyIleIleValLeuAlaGluGlyArgLeuMetCAsnLeu 398
1204 GGCTGGCTACAGAGTCAACCATCTTCTGTTATGTCATGTCATTCACAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACACTGCTCAGCTGACCTCTACGAAAGAAAGA.....GGAATTC 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
415 nGlnValIleAlaGlnLeuGluLeuThrPheGluLysAlaSerGlyLys 432
1298 TCGAGCAAGAGGTTTACACATTCGCAAGCATCTGATGAAAGTCGCT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
432 YGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
1348 GCGCTCCACCTGGATCTCGATGTCACACTTACAAAGCTTACACAGAA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
449 AlalaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLys 465
1398 GCAGGCTGATCATCATCAAGTTCAGTGTGAGGCTCTTACAAAGTCGATG 1447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
465 rGlnSerAspTyrIleSerIleProIleGluGlyProTyrLysLeuArgL 482
1448 CTTACCCTTAT 1458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
482 eUtyrArgTyr 485

```

seq\_name: /cgn2.6/ptcddata/1/paa/US09\_NEM\_COMB.pep:US-09-791-537-23700

```

seq_documentation_block:
; Sequence 23700, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
; SEQ ID NO 23700
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Caltharanthus roseus
US-09-791-537-23700

```

## alignment\_scores:

Quality: 1520.00 Length: 488  
 Ratio: 3.848 Gaps: 6  
 Percent Similarity: 80.943 Percent Identity: 62.295

## alignment\_block:

US-09-759-990-1 x US-09-791-537-23700 ..

Align seg 1/1 to: US-09-791-537-23700 from: 1 to: 485

```

19 ACTGGTGTCCATTCGAGTACAGAAATGCCGACATCACTCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8  ThreSerSerGlyArgGluTyrGlyValLysAspMetSerGlnAlaAsp 24
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGGCCGTGAAGAACTTACCTTGTCTGAGAAAGAAATGCCAGGCTTATG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluLeuGluLeuAlaGluValGluMetProGlyLeuMet 41
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTTCTGTGAGCGTTATTCGCTTGTACAGCCATTGAAGGGTGTGAGATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 eGlyAlaGluAlaGluLeuProSerGlnProPheLysGlyAlaLys 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGTGTCCTCCACATGACAGTCAGACAGCGCTCCATGAGACACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValIleuIleGluThr 74
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CAGAGCTGTGTCGTATGTCATGTCAGATGGGCTTCGCAACATCTTCTTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValAlaGlyTyrPcysSerGlyAsnIlePheSer 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAGATACAGCCGCTGTGCTATGCTGTGCGGCCCAAGGCAAGCACACA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 hGlnGlnHisAlaAlaAlaIleAlaIleAlaArgAspSerAlaIle..... 105
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GAGAAACCCAGCCGGTATCCAGTCTTGCGCTGGAAGGGCAAAACACTGCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaIleTyrPcysGlyGluThrLeuGlu 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATATGTCGAGACATACATACGCGCTCTACATGG...CCAGATGGTC 415
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 nGluTyrTrpTyrPcysThrGluArgAlaLeuAspTrpPcysGlyProaspGly 131
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 AAGGCCACACAGGTTGTCGATGATGATGGTGTGATGCTACACTCTCTATC 465
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 ..GlyProaspLeuIleValAlaAspArgLysAlaThrLeuLeuIle 147
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466 TCCAGAGCGCTC.....GAATTCGAAACAGCCGGTGTGTTC 503
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 HisGluGlyValLysAlaGluGluTyrLysLysAsnGlyAlaLeuP 164
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 AGAGCCAAACAGAGTCGACAACTCGAATACCGCTGCTTCTTCTACAC 553
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 oAspProSerSerThrAspAsnAlaGluPheGlnIleValIleuThrIle 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
554 TCAACGAGGTCTTCAACCAAGACAGACACACGACACAGTTTGCTGCC 603
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 LeuArgAspLysLeuLysSerAspProThrLysTyrThrArgMetLysGlu 197
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
604 GGCATGAACGGTGTTCGAGAGACACAGAGGTGTCACCGCCCTTA 653
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 ArgLeuValGlyValSerGluGluThrThrThrGlyValLysTrpLeuTyr 214
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
654 CCAGCTCGAAGAGAGGCAAACTCTCTTCACAGCCATCAACGTCACAG 703
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 rGluMetGlnAlaAsnGlyThrLeuLeuPheProAlaIleAsnValAsn 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
704 AGCGTGTACAAAGTCGAGTTCGATTAACATACAGCGCTGCCCTCCACTCC 753
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 sPseValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSer 247
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
754 CTTATTCAGTGTATCAACCGCTGCTTCGATGTCATGATGCGCGCAAGAC 803
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

248 LeuProaspLysLeuMetAlaThrAlaThrAspValMetIleAlaGlyLysVal 264
804 AGCTTCGTCATGAGGTTACGGCGATGTCGGAAAGGGCTGCGCTCAATCC 853
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 lAlaValAlaAlaGlyTyrGlyAspValGlyLysGlyCysAlaAlaAla 281
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
854 TCCGTGGCCAAAGCGCTCGGCTTATCATACAGAACTGACCAATCTGC 903
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 euLysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCys 297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
904 GCTCTCCAGGCTCCATGTAAGGCTACAGGTCGCGCCCATCGAAGAGT 953
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298 lAlaLeuGlnAlaThrMetGluGlyLeuGlnAlaLeuThrLeuGluAspVal 314
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
954 CGTCAAGATGTGATATCTTGTACATGACAGCAAGAACTGTCATATCA 1003
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 lValSerGluAlaAspIlePheValThrThrThrGlyAsnLysAspIle 331
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1004 TCTCTGTGACATGAGCCGACATGAAGATGAAGATGAAGCTATTTGCTAAC 1053
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331 lMetValAlaAspHisMetAlaGlyLysMetLysAsnAlaIleValCysAsn 347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1054 ATCGGCACTTCGATACAGAAATGATTAAGATGAGCTTCATGAATACCC 1103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
348 lIleGlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrP 364
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1104 AGGATTCAGACATCCCAATCAAGCAGAAATCAGACATGTGGAAATTC 1153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 oGlyValLysAlaGlyIleThrIleLysProGlnThrAspArgTyrAlaPhe 381
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1154 CAGATGGCCACGCT...ATCTCTCTTGTGTCGAGGCGCCCTTCTAAC 1200
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
381 rOAspThrAsnSerGlyIleIleValLeuAlaGluGlyAlaGlyLeuMetAsn 397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201 CTTGCTGCGCTACAGGTCACCATCTTGTGATATGTCATGTCATTCAC 1250
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 LeuGlyCysAlaThrGlyHisPheSerPheValMetSerCysSerPheTh 414
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1251 AAACGACAGACATGCTCAGCTCAGCTCAGCTCAGTACGAAAGAG...GGA 1294
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 rAsnGlnValAlaIleAlaGlnLeuGlnLeuIleuTrpAsnGluAlaGlyTh 431
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1295 ATCTCGAAGAGAGGTTTACACACTTCGACAGCACTCCAGACATCTCGAAGAGTC 1344
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 yStrGluLysLysValIleValIleuProLysHisLeuAspGluVal 447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1345 GCTGCGCTCCACCTCGGATCTCTGATGTCCACCTTACAAAGCTTACACA 1394
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 lAlaAlaLeuHisLeuGlyLysLeuGlyAlaLysLysLeuThrLys 464
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1395 GAAGCAGGCTGATACATCAACGTCCTGATGAGGCTCTTCAAGTCTG 1444
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
464 sAspGlnAlaAspTyrIleSerValProIleGluLysProTyrLysPro 481
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1445 ATGCTTACCGCTTAT 1458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 lAlaIleTyrArgTyr 485
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq\_name: /cgn2\_6/plodata/1/paa/US09\_NEW\_COMB pep: US-09-935-625-2154

seq\_documentation\_block:

Sequence 2154, Application US/09935625

GENERAL INFORMATION:

APPLICANT: N. ALEXANDROV et al.

TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF

FILE REFERENCE: 2750-1481P

CURRENT APPLICATION NUMBER: US/09/935, 625

NUMBER OF SEQ ID NOS: 33136

SEQ ID NO 2154

LENGTH: 467

TYPE: PRT

ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: peptide  
LOCATION: 1..467  
OTHER INFORMATION: Ceres Seq. ID no. 3439073  
US-09-935-625-2154

alignment\_scores:  
Quality: 1518.00 Length: 470  
Ratio: 3.963 Gaps: 4  
Percent Similarity: 81.489 Percent Identity: 63.404

alignment\_block:  
US-09-759-990-1 x US-09-935-625-2154 ..

Align seg 1/1 to: US-09-935-625-2154 from: 1 to: 467

```

70 GCCCGTAAAGAACTTACCTTGTGAGAAAGAAATGCCAGGTCTTATGCT 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7  GATATGLeuGluLeuGluLeuAlaGluValGluMetProGlyLeuMetAl 23
   : ||| : : : : : : : : : : : : : : : : : : : : :
23 acysatgtrngluhneglpProserGlnProhneylsGlyAlaArgilet 40
   : ||| : : : : : : : : : : : : : : : : : : : : :
120 TCTTGTGAGCGTTATTCGCTTAAAGCCATTGAAGGCTGTCCAGAACTT 169
   : ||| : : : : : : : : : : : : : : : : : : : : :
23 acysatgtrngluhneglpProserGlnProhneylsGlyAlaArgilet 40
   : ||| : : : : : : : : : : : : : : : : : : : : :
170 CTGTTCCCTCCACATGACAGTCCAGACAGCCGCTTCATCGACAGACTC 219
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
40 hrglySerleuHisMetThrIleGlnThrAlaValLeuIleGluThrleu 56
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 AAGAGCTTGGTGTGATGTCAGATGGGCTCTGCAACATCTTCTTAC 269
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
57 ThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerTh 73
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 ACAAGATACAGCCGCTGCTATGTTGTCGCCCAACAGGACACACAG 319
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 rGlnAspHisAlaAlaAlaAlaAlaAlaAlaArgAspSerAlaAla..... 87
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 AGAAGCCAGCCGCTATCCAGCTTGGCTGAGAGGCGGAAACACTGCCA 369
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 .....ValPheAlaTrpLysGlyGluThrleuGln 97
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
370 GAATATGAGAGAACATACACATACCGGCTCTACATGCGCAGATGCTAACG 419
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GluTrpTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGly 114
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 CCCACAGCAGGTTGTCATGATGTCGTCGATGTCATACATCCATCTTCCA 469
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 yProAspLeuIleValAspAspGlyAlaAspAlaThrleuLeuIleHisG 131
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
470 AGGCGCTCGAA.....TTGAAACAGCCGCTGCTGTCCAGAG 507
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 IuGlyValIlysalagluGluIlePheGluIlyThrGlyGlnValProAsp 147
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
508 CCAACAGAGGTGACACACCTGGAATACCGCTGCTTGTGACACTGCA 557
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 ProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLeu 164
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
558 GCAGGTCTTCAACCAAGACAGACACACCTGTCGTCGCGCA 607
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 sGluGlyLeuGlnValAspProLysIlyThrIlyHisLysMetLysGluArg 181
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
608 TGAACGCTGTTCGAGAGACAGACACAGGTGTCCACCGCTCTACAG 657
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 euValIglYValIserGlnGluThrThrGlyAlaLysArgLeuYrGln 197
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 CTCGAGAAGGAGGCAAACTCTCTCCAGCATCAAGCTCAAGCAGC 707
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 MetGlnGlnAsnGlyThrleuLeuPheProAlaIleAsnValAsnAsp 214
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
708 TGTTCACAAGTCCAGTTCGATACATCTTACGCTCCGCTCCCTTA 757
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 rValThrLysSerLysPheAspAsnLeuGlyrGlyCysArgHisSerleuP 231

```

```

758 TCGATGTATCAACCGTGTCCGATGTATGATGGCGGCAAGACACT 807
   ||||| : : : : : : : : : : : : : : : : : : : : :
231 roAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyValAla 247
   : : : : : : : : : : : : : : : : : : : : :
808 CTGTCATGAGGTTCAGCGCATGTGCGCAGAGGCTGCGCTCATCCCTCCG 857
   : : : : : : : : : : : : : : : : : : : : :
248 ValIleCysGlyTyrglyAspValIglYAspGlyCysAlaAlaMetLys 264
   : : : : : : : : : : : : : : : : : : : : :
858 TGGCCAAGCGCTCGGTTTCATCAAGAACTGACCAATCTGCGCTC 907
   : : : : : : : : : : : : : : : : : : : : :
264 sThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 281
   : : : : : : : : : : : : : : : : : : : : :
908 TCCAGGCTGCTGAGAGGCTTACAGCTCCGCGCATCGAGAGACTGCT 957
   : : : : : : : : : : : : : : : : : : : : :
281 euGlnAlaLeuMetGluGlyLeuGlnValIleuThrleuGluAspValVal 297
   : : : : : : : : : : : : : : : : : : : : :
958 AAGATGTGATATCTCTGCTTACATGCAAGAACTGCGATATCATCTC 1007
   : : : : : : : : : : : : : : : : : : : : :
298 SerGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleLeu 314
   : : : : : : : : : : : : : : : : : : : : :
1008 TGTTCACATGATGAGCCAGATGAAGATTAAGCTATTGTGCTAACATCG 1057
   ||||| ||| : : : : : : : : : : : : : : :
314 tValAspHisMetLargLysMetLysAsnAsnAlaIleValCysAsnIleG 331
   : ||||| ||| : : : : : : : : : : : : : : :
1058 GCCACTTCGATTAACGAATTGATACAGATGCGCTTATGAATACCAAGC 1107
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
331 LysIleAspAsnGlnIleAspMetLeuGlyLeuGluThrTrpProGly 347
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1108 ATGCAAGCAATCCCAATACAGCAGATTAAGCAATGAGGAATTCACAGA 1157
   : : : : : : : : : : : : : : : : : : : : :
348 ValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProG 364
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1158 TGGCCAGCT...ATCTCTCTTCTGCTGAGGCGCGCTTCTTAACCTTG 1204
   ||| ||| : : : : : : : : : : : : : : :
364 uThrLysAlaGlyIleIleValLeuAlaGluIlyArgLeuMetLysLeu 381
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1205 GCTGCGCTACAGGTCACCCATCTTCGTTATGCAATGTCATCAACAAC 1254
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
381 LysCysAlaThrGlyHisProSerPheValMetCysSerPheThrAsn 397
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1255 CAGACACTGCTGACCTGACCTCTACGAAAGAGA.....GGAAATCT 1298
   : : : : : : : : : : : : : : : : : : : : :
398 GlnValIleAlaGlnLeuGluIleuThrAsnGluLysAlaSerIlyLys 414
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1299 CGAGAAGAAGTTTACACACTTCCGAAAGCATGTCGATGCAAGTCCGCTC 1348
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 rGluLysLysValIlyValLeuProLysHisLysAspGlyValAlaAl 431
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1349 GCGTCCACCTCGATCTCTGATGTCACCTTACAAAGCTTACACAGAG 1398
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
431 euLeuHisLeuGluLysLeuGlyAlaArgLeuThrLysLeuSerLysAsp 447
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1399 CAGGCTGATACATCAAGCTTCCAGTTGAGGCTTACAAAGTGTATGC 1448
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
448 Gln***AspLysValSerIleProIleGluGlyProThrLysProProH 464
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1449 TTACCGTTAT 1458
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
464 sTyraTrgIlyr 467

```

seq\_name: /cgn2\_6/plodata/1/paa/US09\_NEW\_COMB.pep:US-09-791-537-13836  
seq\_documentation\_block:  
: Sequence 13836, Application US/09791537  
: GENERAL INFORMATION:  
: APPLICANT: Biocomix, Inc.  
: APPLICANT: Debe, Derek  
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY  
: FILE REFERENCE: 261/210  
: CURRENT APPLICATION NUMBER: US/09/791, 537





Mon Aug 19 15:16:20 2002

us-09-759-990-1.n2p.rapn





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:36:01 ; Search time 399.1 Seconds  
(without alignments)  
9569.466 Million cell updates/sec

```

Title:          US-09-759-990-1
Perfect score:  1461
Sequence:       1 atgcttgcacatcactac.....ctgatgttaccgttattaa 1461

```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

```
Searched:      1664314 seqs, 1307041093 residues
Total number of hits satisfying chosen parameters: 3328628
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : Pending_Patents_MA_New.*
1: /cgn2_6/ptodata/1/pna/PCF_NEW.COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW.COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW.COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW.COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW.COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US60_NEW.COMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	614.2	47.0	1428	8	US-60-360-039-11953	Sequence 31953, A
2	547	37.4	1500	5	US-09-540-2098-2680	Sequence 2680, App
3	526.8	36.1	1407	8	US-60-360-039-19432	Sequence 39432, A
4	526.8	36.1	1407	8	US-60-360-039-39816	Sequence 39816, A
5	511	35.0	1483	8	US-60-360-039-44719	Sequence 44719, A
6	508.2	34.8	1427	5	US-09-935-625-2151	Sequence 2151, App
7	503	34.4	1383	8	US-60-360-039-15153	Sequence 35153, A
8	503	34.4	1383	8	US-60-360-039-38339	Sequence 38339, A
9	503	34.4	1392	8	US-60-360-039-38812	Sequence 38812, A
10	497.8	34.1	1338	8	US-60-360-039-39064	Sequence 39064, A
11	492.8	33.7	1377	8	US-60-360-039-34358	Sequence 34358, A
12	481.2	32.9	1398	8	US-60-360-039-32437	Sequence 32437, A
13	479	32.8	1383	8	US-60-360-039-38470	Sequence 28470, A
14	479	32.8	1383	8	US-60-360-039-41230	Sequence 31230, A
15	475.6	32.6	1392	8	US-60-360-039-40463	Sequence 40463, A
16	471.4	32.3	1494	5	US-09-935-625-736	Sequence 736, App
17	470.6	32.2	1389	8	US-60-360-039-44516	Sequence 44516, A
18	462.8	31.7	1401	8	US-60-360-039-35797	Sequence 35797, A
19	458.8	31.4	1395	8	US-60-360-039-11621	Sequence 11621, A
20	452.8	31.0	1404	8	US-60-360-039-32928	Sequence 32928, A
21	452.8	31.0	1404	8	US-60-360-039-33073	Sequence 33073, A
22	448.6	30.7	1348	8	US-60-360-039-37862	Sequence 37862, A
23	441.4	30.2	1398	8	US-60-360-039-43078	Sequence 43078, A
24	434.6	29.7	1341	8	US-60-360-039-41239	Sequence 41239, A
25	430	29.4	1389	8	US-60-360-039-43361	Sequence 43361, A

26	428.8	29.3	1389	8	US-60-360-039-31609	Sequence 31609, A
27	355.8	24.4	1341	8	US-60-360-039-27370	Sequence 27370, A
28	331	22.7	2429	7	US-10-116-802-98	Sequence 98, Appl
29	331	22.7	2658	5	US-09-785-276A-23115	Sequence 23115, A
30	331	22.7	2658	5	US-09-785-276A-28985	Sequence 28985, A
31	302.4	20.7	1299	8	US-60-360-039-26112	Sequence 26112, A
32	302.4	20.7	1299	8	US-60-360-039-26117	Sequence 26117, A
33	293.4	20.1	1398	5	US-09-762-154-51	Sequence 51, Appl
34	293	20.1	1284	8	US-60-360-039-34132	Sequence 34132, A
35	265.2	18.2	1278	8	US-60-360-039-33822	Sequence 33822, A
36	265.2	18.2	690	8	US-60-360-039-33976	Sequence 33976, A
37	264.4	18.1	687	8	US-60-360-039-33980	Sequence 33980, A
38	264.4	18.1	687	8	US-60-360-039-33945	Sequence 33945, A
39	250.2	17.1	1350	8	US-60-360-039-25427	Sequence 25427, A
40	244.6	16.7	1356	8	US-60-360-039-35798	Sequence 35798, A
41	244	16.7	3634	5	US-09-785-276A-23184	Sequence 23184, A
42	244	16.3	3634	5	US-09-785-276A-29050	Sequence 29050, A
43	238.4	16.3	874	7	US-10-027-653-151874	Sequence 151874, A
44	238.4	16.3	874	7	US-10-027-653-151875	Sequence 151875, A
45	229.8	15.7	1450	8	US-60-360-039-28784	Sequence 28784, A

## ALIGNMENTS

```

RESULT      1
US-60-360-039-31953
: Sequence 31953, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Xiongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 31953
:
: LENGTH: 1428
:
: TYPE: DNA
:
: ORGANISM: Thermobifida fusca
: US-60-360-039-31953

```

Query Match	Similarity	42.0%	Score 61.4	2	DB 8	Length 1428
Best Local	Similarity	64.4%	Pred. NO. 3	3e-171		
Matches	919	Conservative	0	Mismatches	508	Indels
						Gaps
						0
QY	31	ttcagatcacgaatttcgacatcacatccatggtctcgcgcgttaaggaacttaccctt	90			
Db	1	ttcgacttcaagtcgcgcacacctcgtcgtcgcgaagttcgcgtcgaagagatccgcttc	60			
QY	91	gctcgagaagaataagccaaggtcttaagtgtcttcgtgacgcgttatctccgcttcaagcca	150			
Db	61	gccgaacaacagagatgcgcggctctatgaccaccgcggcagaaatccgaaanaaacgcg	120			
QY	151	ttgaagaagttcagaaatctcgttccctccacatctacagttccagacagccgtctccatc	210			
Db	121	cttcgggggtgcgaagatcacagcgtcgcgtcgcaaatgacccgttcaaacgcgctcgtatc	180			
QY	211	gagacactcacagctctctgtgtcgtatgtcagatggtccttcctcgcaacatcttctcca	270			
Db	181	gagacctgtgtgagttcttggtgcgttgaatctcgtctggcgcagctcgcaacatcttccacc	240			
QY	271	caagatatacagccgcgtcgtcgtatcgttctcggcccaacagagccacacagagaagccagcc	330			
Db	241	caggacaacagccgcgcgcgcgt	300			
QY	331	ggatcccagctcttcgctcttgaaaggcgaaacactcccaacatacttggagaacacatac	390			

Db 301 ggggtgctggtgtcgtgtaagggcgaaacgctcgaagaagtaagtgtgtgtgcaaggaa 360  
Oy 391 cggcgtctacatggtgcagatgtgtcaagggccacagaggtgttcgagatgtgtat 450  
Db 361 caggcgctcaactggtgcggagcgcggaagggcccgaaatgatacccgaaagggcggaac 420  
Oy 451 gctacacgtctcactcctcaaggggttcgaattcgaaacacgggtgtgttccagagca 510  
Db 421 gcaacacgtgtgttccaaaggggtgtagtaagaaggggggtgcgttcccgaccgg 480  
Oy 511 acaagaagtcgaacactcgaataccggtgtgttctgttaacactcgaacaggtttcaac 570  
Db 481 gataccgcgaagcagaggttccgcatcgtctgtgaactgtcccgaggaacgttggcg 540  
Oy 571 caagacaagaacacactggtgcacacagttgtcgtcgccgcatgaaacgtgttccgaagaca 630  
Db 541 gaaacaccccgcaagtggaagagatcgtctccggatcaaggggtgtacccgaggaagac 600  
Oy 631 acaacagtggtccacacgctctcaacagctcgagaagaggcaaatcctctcccgacc 690  
Db 601 acgacgggggtgacacgctctcagagatgcaaggttccgacagctgttcccgcg 660  
Oy 691 atcaaggtcaacgacgctgttacaagttccaagtgcgaataacatcagcgtcgccgcaac 750  
Db 661 atcaacgttcaacgactccgctcgaagaagagagttcgaacaagttacggtcgccgcaac 720  
Oy 751 tcccttatcgatgtatcaacacgctgtcctcgatgtatcagtcgagcgagcaacagctctc 810  
Db 721 tgcgtatcgagcgatcaacacggcgccatgagctgtcctcagtcgagcgagcgagtg 780  
Oy 811 gtcaatggttaagcgatgtcgcgaaggggtcgctcaatcctccgtgtgccaagcgct 870  
Db 781 gtggcggtgatacggcagcgtcgttaaggtcgcgagtgctcgcgagcgcaagggcgcc 840  
Oy 871 cgcgtatatacgaagaactcgaacccacatcgtcgtcctcccgagctgcgcatgaaaggtac 930  
Db 841 cgcgtatcgtcgaagatcgacacgagatcgtcgtcgtcgaacggcgcatgtgagcggtac 900  
Oy 931 caggtcgcgcgacgcgaagagtcgtcaagagttgcaatcctcgtgtgccaagcgct 990  
Db 901 caggtgacacgctcgtcgaaggtgtcgtcgaacggcgcatgttcatcaccgacggcg 960  
Oy 991 aactgtgatacctctcgtgtgaatgtgtccagatgaaaggtatgtgtgt 1050  
Db 961 aacccggagcgtcattactgtcgagacatgtgcccgaatgaaaggtatgtgtgt 1020  
Oy 1051 aacatcgccactcgaatacgaatgtatcaagatgtcccatgaaatccagacatc 1110  
Db 1021 aacatcgccactcgaatacgaatgtatcgaatgtggtggtcgccgcatcccggtatc 1080  
Oy 1111 aagcacatcccaatcaagcgaatagatgtggaattcccaatgtgccaagctatc 1170  
Db 1081 gagaagatcgagatcaagcgcgaggtgtcgaaggtgagcgttcccgagcggtatccatc 1140  
Oy 1171 ctcccttctgtgtgagggcgcttcttaactgtgtggtcggttaacgttcaacccatcttc 1230  
Db 1141 atcgtgtcgtgtgagggcgctgtcgaactgtggaactgtggaacggcgaccccgagcttc 1200  
Oy 1231 gttatgtcaatgtcattcaacaacacgaacactcgtcgaactgtggaactgtcaagaagaga 1290  
Db 1201 gttatgtcgaacacgcttcaataacaggtcatcgtcgcaagctgtgttcccaagacc 1260  
Oy 1291 ggaacatcgaagaagaaggtttacacacttccgaagatcgtatgaaaggtcgtcgc 1350  
Db 1261 gaccagttaccgacggcggtgtacgtctgtcccaagcactgacgagaaggtcccgccg 1320  
Oy 1351 ctccacactcggtactcgtgtgtccacacttcaagaaggttaacagaagcaggtctactac 1410  
Db 1321 ctgcactgtgacgactcgtgtgaaactgtacccggtccacaaaggaacagcgccgaglac 1380  
Oy 1411 atcaacgttccagttgaggggttccattcaaggtctgattacgtta 1457  
Db 1381 atcggaggtggaatcgaacggcgccctacaaagtccgaacactacgtta 1427

RESULT 2  
US-09-540-209b-2680  
; Sequence 2680, Application US/09540209B  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOID  
; FILE REFERENCE: 2709, 1001-001  
; CURRENT APPLICATION NUMBER: US/09/540, 209B  
; NUMBER OF SEQ ID NOS: 10444  
; SEQ ID NO 2680  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: B. fragilis  
US-09-540-209b-2680  
  
Query Match 37.4%; Score 547; DB 5; Length 1500;  
Best Local Similarity 63.4%; Pred. No. 2.5e-151;  
Matches 904; Conservative 0; Mismatches 485; Indels 36; Gaps 3;  
  
Oy 37 tacagaattgcgacatcaactcactatgttctcgcgcgtaaggaactaccctgtgag 96  
Db 112 tacaagtggagacatcacgttgcgacttcggtcgcgaagaaatcgaactgtgcagaa 171  
Oy 97 aagaagaatgcaggtcttatgttctgtgtgagcggtatctccgtcttaagcattgaa 156  
Db 172 aagaagaatgcggtcgttaatgtccttcgcgaaataatgagaacccaacattgaa 221  
Oy 157 gttcgaagaatcgtgttccctccacatgaaatgcagacgacgcgttccatcagaca 216  
Db 232 ggtgcgggtatcatgaggtacactatgacatgataccagacgcgcgttccatcagaca 291  
Oy 217 ctcaagctctgtgtgtgtatgtcagatgtggtctccgaacatctctcacaagaat 276  
Db 292 ctgttgacccctgggtgaggtgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 351  
Oy 277 acaacgcgtgtgtatgt 336  
Db 352 catgtctcagcagctatcgcag -----ctcaggtgt 384  
Oy 337 ccaattcgcgttgaagggcgaaacactcccaatactgtggagaacacatccgcgt 396  
Db 385 gccatttgcgttgaagggaaacatgtggaattatgtgtgtgtgtgtgtgtgtgtgtgt 444  
Oy 397 ctcaatggtcagatgttcaagggccacagcaggtgtgtgtgtgtgtgtgtgtgtgtgt 456  
Db 445 ctgaacttgaagggggaaggaacagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 504  
Oy 457 ctccctatcccaaggggttcgaattcgaacacgcggtgtgtgtgtgtgtgtgtgtgtgt 516  
Db 505 atgatgttcatatgtatgtatgagcagaagaacaaatgtcgtgtatgtgtgtgtgtgtgt 564  
Oy 517 gctgaacactcgaataacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 576  
Db 565 catgctcgaagcagaatagaaat---tgaatgtatctgaataaagtgctgtgcagaagac 621  
Oy 577 aagaacacatggtgcacacagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 636  
Db 622 aagaagaatgt 681  
Oy 637 ggtgtcgaacgcttcaacagctgtgagaagggcgaacactcctctccacacatcaac 696  
Db 682 ggtgtcatcgtccatacacaatgtcagaagaaggaaggtgtgtgtgtgtgtgtgtgtgt 741  
Oy 697 gtcaacacgctgttacaagaatltcaagttcgaatacattcagcgtcgcgcacatccct 756  
Db 742 gtaaacgactcgtgttacaacaaatcgaacactcgtgtgtgtgtgtgtgtgtgtgtgt 801  
Oy 757 atcgaatgtatcaacgctgt 816

Query Match	36.1%	Score 526.8;	DB 8;	Length 1407;
Best Local Similarity	62.5%;	Pred. No. 2.4e-145;		
Matches 897;	Conservative 0;	Mismatches 497;	Indels 42;	Gaps 3

QY	34	gggttaagaattgcgcacatcaacaactccattctctctcttcggttaaggaaacttaaccttgc	93
Db	1	gattacaagaatcgccgaacatctcccttcggccgattggggccgcaaggagcttgagactatgc	60
QY	94	gggaagagaataatccagagctcttaattgattctctctgtagaggttaattccgcttcaagactg	15
Db	61	ggacgcagagattgcgggcttgatctgatactccgcgaagaaatgagcaagcaagccgctg	12
QY	154	aagggtgtcagaagattctctgttccctccataatgacagttccagagaccggtccatcagag	21
Db	121	aaagacgttcgcatactcaaccggcttcgcgcagaaatgacatccatccagaccggtgctatcgaa	18
QY	214	acaactcaagactcttctgtgtctgataatgagatgggtcttcgcgaacatctctctcaaca	27
Db	181	acgctcaagagacatcgcgcgaattggtgcgttcggcttcgcgaacatctctccaccag	24
QY	274	gatacagcgcgtgcgtctatctgtctgcgcaccaagcaaccagaagaagccagccggt	33
Db	241	gaccacggcgccgcgcgagatttgcg-----ccaccgcg	27
QY	334	atccacagcttctgccttggaaggcggaacaactccagaatactggaggaacacataccgc	39
Db	274	acgcgcgggtgttcgcttggaaggcggaagacgcgtggaagaattatgtgagcttgcacctgac	33
QY	394	gcttccaaatgtgcagatggtca-----agcccaacagcaggtgtctgaatg	44
Db	334	gcgcttgacatttcaaccttcgcccagcagcaaccttgacccgcgcgcgagctgtgtgtgagac	39
QY	442	ggtgtgtagtgcatacctctcaatcccaagggtcttcgaattccgaacaacgcggtgctgt	50
Db	394	ggcgcgagacgttcaaccttctgataccacaagggtctgtgactgcaaaacggaacacactg	45
QY	502	ccaagagccaacagaagcttgacaacacctgcaatccgactgcgtgtcttcttcaacatcaagcag	56
Db	454	gtcgaacgaagc---gacctctcgaacgaagaaggcgtatataaggcaactgtctcaagcgc	51
QY	562	gtcttcaaccagaagaagaacacctgtgcacacagttgtctgcgcgcatgaaacggtgttc	62
Db	511	gttgctctgcagcgtcccgaggtaactggcgcggtgtgtgataagaaatgaaaggcgtctct	57
QY	622	gaaagagaacaacaagggtgtccaacgcgctcttcaacggctccgacgaaagaggaagcaacctcc	68
Db	571	gaaagagacccaacacggcggtgcacgcgctgtatccagatgtccgaaagccgcgaacctgcg	63
QY	682	ttccccaagccaatacgttcaacagacgcgtgttcaaaagttccaaagtctcgataacatactacgcg	74
Db	631	atcccgagccaatacgttcaacagactctgcatacaagagcaagttctgcacaaccttaagcc	69
QY	742	tgcgcacaactcccttaatcgatgtgatacaacgcgttcccgatgtatagatccggcggaag	80
Db	691	tgcgcagagtcgcttgccgagatgccttcaagcgcgcgataagcgtatgactgtgtgcgcgcaag	75
QY	802	acaagctctgcataatgggttaacggcgaatgtccgcgaagggctgcgcatactccctctggc	86
Db	751	gtccggtgtgtctgcggttaacggcgaatgtccgcgaagggcgcgcgcgcgcgtcgtgcgtcc	81
QY	862	caagagggcccggttatcatcaacagaactcgaccacaatctgcgctctccaggtctgcag	92
Db	811	tacggtgcccgcgtcatctgcatacccgaaatagacccgactctgcgccttcgacggtcgtgt	87
QY	922	gaaagctacaggttccgcgcgcatacgaaggaaagtcgtcaaaagatgctcgataatcttgta	98
Db	871	gaaagcttcgaggttcaacaaccatcgataccaaccttgccgcgcgcgcgcgcataatgtatgc	93
QY	982	tgcagaaggaatcgcgataatcatctctgtgtgacatgataggcccaagatgaaagataagct	104
Db	931	accacccgggaacaaggaacatacataccggtcgaagcaacctggaacgcatgaaagaccagcc	99
QY	1042	atgtcgttcaactcgcgcacttgcataaagaaattgatacagatggccctcagaataac	110
Db	991	atcgcttgcacaatcgcgcacttgcgaacaacgcgagatccaggtcgatgtgcgtgaatgacct	105



RESULT 5  
US-60-360-039-44719  
; Sequence 44719, Application US/60360039  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Chen, Xianfeng  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)A  
; CURRENT APPLICATION NUMBER: US/60/360,039  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 44719  
; LENGTH: 1383  
; TYPE: DNA  
; ORGANISM: SYNECHOCOCUS SP. WH 8102  
US-60-360-039-44719

Query Match 35.0%; Score 511; DB 8; Length 1383;  
Best Local Similarity 62.0%; Pred No. 1.1e-140;  
Matches 883; Conservative 0; Mismatches 500; Indels 42; Gaps 3;

QY 34 gagtacaagtgcgcacatacctcatgtctctgcgcgttaagaactacccttgct 93  
DB 1 gattcgctattgctgcacatacctcatgtctctgcgcgttaagaactacccttgct 60  
QY 94 gagagaagaatgctgcgtctctctctctctctctctctctctctctctctctct 153  
DB 61 gagacgaagaatgctgcgtctctctctctctctctctctctctctctctctctct 120  
QY 154 aagagtgctgacgaatct 213  
DB 121 aagagtgctgacgaatct 180  
QY 214 acactacagctctgt 273  
DB 181 accctgt 240  
QY 274 gatacagcgtctgt 333  
DB 241 gatacagcgtctgt 273  
QY 334 atcccaagctctgcgtctgcgaagagcgaaacatcccaagctctgcgaagagcg 393  
DB 274 atcccaagctctgcgtctgcgaagagcgaaacatcccaagctctgcgaagagcg 333  
QY 394 gctctacatgctgcgaagagcgaaacatcccaagctctgcgaagagcg 453  
DB 334 atcccaagctctgcgtctgcgaagagcgaaacatcccaagctctgcgaagagcg 393  
QY 454 acactcctcatctccaaagagctgtgaattcgaaacagcggtgtgtgtgtgtgtgt 513  
DB 394 acggt 447  
QY 514 gaaagctgacaaactgcgaagagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 573  
DB 448 aaccctcccaag 507  
QY 574 gacagaagaacactgacaaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 633  
DB 508 gacccaacact 567  
QY 634 acaagtgctacacgctctacacgctctacacgctctacacgctctacacgctctac 693  
DB 568 aacggt 627  
QY 694 aacgtaacagagctgt 753

DB 628 aacgtcaacgactctgcacaaagagcaagctgcacaaactctacgctgcgcgagctg 687  
QY 754 ctatcgatgtatcaacacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 813  
DB 688 ctgtgtgacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaac 747  
QY 814 atggtatcgagctgt 873  
DB 748 atggtatcgagctgt 807  
QY 874 gtatcatcaacgaactgcgaacacatctgcctctcaggtgtgtgtgtgtgtgtgt 933  
DB 808 gtctgcacgcgaag 867  
QY 934 gtccgc 993  
DB 868 gt 927  
QY 994 tgcgaatacctctgt 1053  
DB 928 taccaggt 987  
QY 1054 atcgccactctgt 1113  
DB 988 atcgccactctgt 1038  
QY 1114 caatcccaatcaacgaacgaacgaacgaacgaacgaacgaacgaacgaacgaacga 1173  
DB 1039 tgggaacactcaacgcccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1098  
QY 1174 ctctctgt 1233  
DB 1099 ctgt 1158  
QY 1234 atgtcatgtcatcaacaaacgaacgaacgaacgaacgaacgaacgaacgaacga 1293  
DB 1159 atgaacaaactctcaacaaacgaacgaacgaacgaacgaacgaacgaacgaacga 1218  
QY 1294 aatctcgagaagaaggtttcaacactctccgaacgaacgaacgaacgaacgaacga 1353  
DB 1219 gagtaaggaagaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1278  
QY 1354 caactgt 1413  
DB 1279 caactgt 1338  
QY 1414 aacgttcaggt 1458  
DB 1339 aacgttcaggt 1383

RESULT 6  
US-09-935-625-2151  
; Sequence 2151, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 2151  
; LENGTH: 1727  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: Misc\_feature  
; LOCATION: 1..1727  
; OTHER INFORMATION: Ceres Seq. ID no. 3439070  
US-09-935-625-2151

Query Match	34.8%	Score 508.2;	DB 5;	Length 1727;
Best Local Similarity	62.0%	Pred. 8.1e-140;		
Matches 898;	Conservative	0;	Mismatches 503;	Indels 48;
				Gaps 4

OY	34	gagtcacgaattgycgcagacacaaacccacacatgctctcgcgcgttaagaagaaattacacgt	93
Db	104	gataccaaggtcaaaagacactgtctcaagccgatttcgtctcgtctcgaaccgcgcgtccgc	163
OY	94	gagaaagaaatgycagagttctatggtcttcttcgtgagcggtatctccgcttctaagccattg	153
Db	164	gaagttcgaatgctcgtggaactcagctgtcttcgtgacccaatttcggactcttcagccattc	223
OY	154	aaggtgtcagaatctctgttccctccagatgacagtcgcagacgcgtccctcattcgag	213
Db	224	aaagcgctcagatcacatccggaactctctcaatgatcacatccaacccgcgttactatcgaa	283
OY	214	acaactcaagctctctggtgctgatagtcagatgaggtcttcctgtgcaacatctctctaca	273
Db	284	accctaactgcctctcgtgtctgaagtcacatggtgtgtctcgtgcaacatctctccacac	343
OY	274	gatacagccgcgtcgtctatcgtgttcgtgcgcgcacacagcgacacacagagaagccgcgt	333
Db	344	gacacacgcgcgcgcagcatctcgtctgactccgc-----	378
OY	334	atcccaagtcctcgcgtgagaagggcgaaacacatcccaagaatactgagggaacacataccgc	393
Db	379	--cgctgtttctcgcctggaagaaagtgtgagactcttcacagagatctcgtgtglaaccagcgt	436
OY	394	gctctcaacatgycacagatgctcgaagccccaacagctgtctcgatgatactgtgtgtaagct	453
Db	437	gctctcagatgtgggtctccaggtggtgttcctcgatctgattgttgatgctgtgtgagcgt	496
OY	454	acactccctatctccaaggtctcgaa-----tctgaaacagccgcgtgctgct	501
Db	497	actcttttattctatcagtaggggtgttaagctgtagagatactttagaagaactgtgcgaagt	556
OY	502	ccagagcccaacagaaactgcacaaacccctcgaaataccgcgtctctcgtctacaataacag	561
Db	557	ccgtgactcactctctactgataaacccctgagtttcagatcgtgtgtctatatacaagaa	616
OY	562	gtctctcaaccaagaacaaagcaacacacagctgtgctgcgcgcgtcagatgaaacggtgttcc	621
Db	617	gttcttcaagttgtgattctctaagaagttcccaagaatgaagagagaagactgtgtgtgtctct	676
OY	622	ggaagagacaacaacaggtgtccacccgcctctacacagctcgcagagaaggggcacaactcctc	681
Db	677	ggggaacactcacacactggtgtgttaagagcttaccaggtgacgaagaaatggaaccttctg	736
OY	682	ttcccaagcacaacagttcaacgacgctgtgttaacaagtccaagtctgataacatactaacgc	741
Db	737	ttccctgcacattacgtttaacgactcctgtctacccaagagcaagttcgcacaactgttatggt	796
OY	742	tgcgcacatcccttactcgatgttatcagaacgcgtctccagatgcatgatactgcgcgcgaag	801
Db	797	tgcgcgtacataactccctgatagtgtccatcagtaggggcacatgatactgatactgcgcgaag	856
OY	802	acagctctcgtatggtgttaacggcgatgctgcgcaagggctcgcgtcaatccctccgtggc	861
Db	857	gttcgtcttatctcgttgataatgttgatgttgtagaagaggtgtgtctgtctgcacatgaagact	916
OY	862	caagggcgcgcgttatcatcacaagaactgcgaacaaatcgcgcctcccaagcgcgcgcacatg	921
Db	917	gttcgtgtcagagatcaatttgtaacttgatgatacccatccatcgtgcctccaagcttgatg	976
OY	922	gaagcttaccaggttccgcgcgcacatcgaggaagatcgtccaagatgctgatalcttcgttaca	981
Db	977	gaagacttcaagttcttcaaccttgagtagagttgtctccaagaagctgatatcttgttcaacc	1036
OY	982	tgcacaggaacatgcgatatcatctcgtttgtagacatgtagcccaagatgaagatcaagct	1041
Db	1037	accacccggttaacaagaacacatcatcgtgtcgacacacatgtaggaagtgaagaaacacgcgt	1096

OY	1042	atgtgcgtgaacatcgcgcattcgataacgaaattatatacagaatggcctcatgaataac	1101
Db	1097	attgtgtgcaacattgtgtcaatttgcataatgaaattgaaactgtgtgaacttggaaactac	1156
OY	1102	ccaagcatcaagaacatcccaatcaagccgaatgaacatgttggaaattcccaatgagc	1161
Db	1157	cctgtgtgtgaagcgtatccatccatcaagaagccacagactcaggttgggtgttcacagaagcc	1216
OY	1162	caagcgt---atctctcttctgtgttggaggcgccctcttaaccttggctcgcgaaggt	1218
Db	1217	aagcgctggaatcaattgtcctgtgcgaagggcgcgtcgtatgaacttgggttgtgcacggct	1276
OY	1219	caaccatctctcgttatgttcaatgtaatacaacaacagacaactcgtcctcagctcagacctc	1278
Db	1277	caaccgaatttcgttgttgccttctcttcttaaccacaacaggtgatgtccagctcagacctc	1366
OY	1279	t-----acgaagaagagagaaatctcagagaagaagtttaacaacttcogaagacatctc	1332
Db	1337	tggaaagaaagaaacaaagcggaaatgcgaagaagaagttgtacgttcttcccaagcatlttg	1366
OY	1333	gatgaagaagtcgctcgccctccacctcagatctctcgtatgttccaccttacaagaacttaca	1392
Db	1397	gatgagaaggttgcattactctcactctgggaagccttggagccaaagcttacaagaactgtca	1456
OY	1393	caagaacagagctgtactacatcaacgltccagcttggaaaggtcccttaacgaactcgtatgctaac	1457
Db	1457	aaggaaccaawctactacgtcagcatctccaattgaagagacatatacaagaagcctctcaatcac	1516
OY	1453	cgttatata 1461	
Db	1517	aggtactga 1525	

```

RESULT 7
US-60-360-039-35153
: Sequence 35153, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 35153
: LENGTH: 1383
: TYPE: DNA
: ORGANISM: Agrobacterium tumefaciens
US-60-360-039-35153

```

Query Match	34.4%	Score 503;	DB 8;	Length 1383;
Best Local Similarity	61.6%;	Pred. No. 2.6e-158;		
Matches 878;	Conservative	0;	Mismatches 505;	Indels 42;
				Gaps 3

[illegible]













```

:
: CURRENT FILING DATE: 2002-02-21
:
: NUMBER OF SEQ ID NOS: 47374
:
: SEQ ID NO 32437
:      LENGTH: 1398
:      TYPE: DNA
:
: ORGANISM: Ralstonia metallidurans
:
: US-60-360-039-32437

```

Query Match	32.9%	Score	481.2;	DB	;	Length	1398;
Best Local Similarity	61.1%	Pred. NO.	7.4e-132;				
Matches	874;	Conservative	0;	Mismatches	508;	Indels	48;
						Gaps	4;

QY	34	gagatcaagaattcgccagacataccaactccatctggtctctgcgcgctaaagaaactaaccttgct	93
Db	10	gactacatactggtgcgcgaataatccgcgcctgcgcgcgttcggccgcgcgaaggaataatccgcgcgc	69
QY	94	gagaaagaaatgcccagagctcttaatgtctctctctctctctctctctctctctctctctctctct	15
Db	70	gagacacgaataatgcgcgcgtctgattgctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	125
QY	154	aagggtgtctcaaatctctcgtgtctctctccatctgaatgacagttccagaacagccgtctcatcag	21
Db	130	aagggc	189
QY	214	acacatcaacagct	273
Db	190	acgtctcaagagcact	249
QY	274	gatacaacgcgcgtctgctctatcgtttgtctggcccaagaagcacacacgaagaagccagccggt	33
Db	250	gatacgc	282
QY	334	atcccagct	39
Db	283	acgcgcgcgtct	342
QY	394	gctctcaatgcccagatggtctcaagggcccaacagcaggtgtctgatgatgtgtgtgtgtgt	453
Db	343	atctctgcacatggtgc	402
QY	454	acacatctcatctctcaagggcgtctctgaattctcaaaaagccgcgcgcgtctctctcaagccaca	513
Db	403	acgcgtctgctgcacccctgc	456
QY	514	gaaagctgcgaacaactctgaataatccgcgtctctctctctctctctctctctctctctctctctct	573
Db	457	aagccgc	516
QY	574	gacaagaacaacccacgc	633
Db	517	gattctgaactctgttacaagc	576
QY	634	acaaagtctccacccgc	693
Db	577	acccgc	636
QY	694	aacgcgtcaacgaacgcgtttaaagaagttccaagtctgcataacatctacgcgcgcgcgcgcgcgc	753
Db	637	aacgcgtcaacgaacgc	696
QY	754	cttatacgtatgtatacaacccgtct	813
Db	697	ctgtgtgacgc	756
QY	814	atgggttttaaggc	873
Db	757	gcccgcgttaacgc	816
QY	874	gtatcatcatcaagaatccatcccaatct	933
Db	817	gtgtgtgtgtacccgaagatctcgatcttggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	876

QY	934	gtccgcgcgacttcgagaaagctcgtcaagaatgctcgatatcttcgttaacatgcacaggaaac	993
Db	877	gtggtgacgattgactacgcgcgcgcgaacaggtgtacacattctcgttaccctgcgcgcgcaat	936
QY	994	tgcgatatcatctctgtttgacatgatgtgcccagatgtgaagagataaaggtcatatttcgtaac	105
Db	937	taccacgtgtgatccgcacagacacacatctgcccagaatgtgaagccagcgcatctgtccgaac	996
QY	1054	atcgcgcacacttcgataacgaatattgatatacagatgtgccttcataagaaataccagcatcaag	111
Db	997	attggcgcactctgcacaacgcaaatctgcatactgcatactgcgcgaagaagtacgaattgtgacgag	1054
QY	1114	caacatcccaatcaagcccgagaatactgcacatgtgtggaattcccaagatgtgcacgcatctctc	1177
Db	1057	-----atcaagccgcgactgttgacacacgctcaagtctccgcgcgcgcgaagaataatcatc	1107
QY	1174	cttctgttcgagagccgcctcttaacacttgctgcgtgcgcgtcaagatcaaccactcttgct	123
Db	1108	atctcgtgcacaagggccgcctcgtgtgaaccttggctgtgtgccaaaggccacacccgtctgtaactg	1167
QY	1234	atgtcaatgtcatctacacaaacccagacacatcgtctacgtcgcaccccttaacgaagaagagga	129
Db	1168	atggagcagactcgtctcgcacaacagacgatactgcacagatcgccagatctgaactgtggccgaacgtgac	1227
QY	1294	aattctcgagaagaaggtt-----tacacacttcgcgaagatctctcgcgatgaagaagtcgct	1347
Db	1228	agcgcgcaagtaaccggtttggtgcgtgtaacgtgtcgtccgaagaagatcttgcgcgaagaaggtcgcg	1287
QY	1348	cgctccacacttcgagatctcctcgtattctcacacttacaagaagcttacaacgaagaagctcag	1407
Db	1288	cgctctgcagctccggaagctgtgacgcgcgaactctgacacgcgactgtacgcgcagcagcgcgcc	1344
QY	1408	tacatcaaatgtccagttgagaggttccttaataagtcgtatgattacggtta	1457
Db	1348	tacatcgcggtgtaagaagaagaagcccggtatacaagccgcacacctaacggtcta	1397

```

RESULTS      13
US-60-360-039-28470
: Sequence 28470, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-1052052A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ. ID NOS: 47374
: SEQ. ID NO 28470
: LENGTH: 1383
: TYPE: DNA
: ORGANISM: Burkholderia fungorum
: US-60-360-039-28470

```

[illegible]

[illegible][illegible]

```

RESULT 14
US-60-360-039-31230
: Sequence 31230, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ. ID NOS: 47374
: SEQ. ID NO. 31230
: LENGTH: 1383
: TYPE: DNA
: ORGANISM: Burkholderia cepacia
US-60-360-039-31230

```

[illegible]





1	1661	100.0	1461	29	US-09-759-990-1	Sequence 1, App1
2	596.4	40.8	1357	23	US-09-606-760A-219	Sequence 219, App1
3	596.4	40.8	1557	23	US-09-746-640A-97	Sequence 97, App1
4	594	40.7	1422	29	US-09-738-628-836	Sequence 836, App1
5	536	36.7	1902	25	US-09-654-617-386668	Sequence 386668, App1
6	536	36.6	1902	27	US-09-684-016-386668	Sequence 386668, App1
7	535.4	36.6	1488	28	US-09-415-888-23	Sequence 23, App1
8	531.6	36.4	1434	28	US-09-703-708-7943	Sequence 7943, App1
9	531.6	36.4	1434	57	US-60-163-320-7943	Sequence 7943, App1
10	531.6	36.4	1434	57	US-60-164-320-651	Sequence 7943, App1
11	531.6	36.4	1434	57	US-60-163-791-7943	Sequence 7943, App1
12	531.6	36.4	1434	57	US-60-163-791-7943	Sequence 7943, App1
13	531.6	36.4	1434	57	US-60-163-791-7943	Sequence 7943, App1
14	531.6	36.4	1434	57	US-60-163-791-7943	Sequence 7943, App1
15	529.6	36.2	1396	23	US-09-606-740A-223	Sequence 691, App1
16	512	35.0	1815	29	US-09-746-660A-101	Sequence 691, App1
17	512	35.0	1815	70	US-60-312-544-2389	Sequence 2389, App1
18	512	35.0	1937	27	US-09-654-617-260094	Sequence 260094, App1
19	508.8	34.8	1865	28	US-09-684-016-260094	Sequence 260094, App1
20	506.6	34.7	1458	28	US-09-708-427-77518	Sequence 77518, App1
21	506.6	34.7	1458	75	US-09-708-427-77518	Sequence 22303, App1
22	506.6	34.7	1458	75	US-60-361-294-15	Sequence 118096, App1
23	506.6	34.7	1969	25	US-09-634-617-118096	Sequence 118096, App1
24	506.6	34.7	1969	25	US-09-684-016-118096	Sequence 118096, App1
25	506	34.6	1401	19	US-09-514-000-5922	Sequence 5922, App1
26	506	34.6	1401	19	US-09-739-449-5970	Sequence 5970, App1
27	506	34.6	1401	31	US-09-803-110-5970	Sequence 5970, App1
28	506	34.6	25572	19	US-09-514-000-416	Sequence 346, App1
29	506	34.6	374732	29	US-09-739-449-416	Sequence 216, App1
30	497.8	34.1	1728	31	US-09-803-110-216	Sequence 216, App1
31	496.8	34.0	1728	22	US-09-620-994B-3581	Sequence 3581, App1
					US-09-565-309A-47857	Sequence 47857, App1





Db 1441 tctgattctaccgttataa 1461

## RESULT 2

US-09-606-740A-219

Sequence 219, Application US/09606740A  
GENERAL INFORMATION:  
APPLICANT: Pompeius, Markus  
APPLICANT: Krieger, Burkhard  
APPLICANT: Schroder, Hartwig  
APPLICANT: Zelder, Oskar  
APPLICANT: Haberhauer, Gregor  
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
FILE REFERENCE: BGI-121CP  
CURRENT APPLICATION NUMBER: US/09/606,740A  
CURRENT FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: 60/141,031  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 60/142,101  
PRIOR FILING DATE: 1999-07-02  
PRIOR APPLICATION NUMBER: 60/148,613  
PRIOR FILING DATE: 1999-08-12  
PRIOR APPLICATION NUMBER: 60/187,970  
PRIOR FILING DATE: 2000-03-09  
PRIOR APPLICATION NUMBER: DE 19930476.9  
PRIOR FILING DATE: 1999-07-01  
PRIOR APPLICATION NUMBER: DE 19931415.2  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931418.7  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931419.5  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931420.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931424.1  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931428.4  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931434.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931435.7  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931443.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931453.5  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931457.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931465.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931478.0  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931510.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931541.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931573.6  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931592.2  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931632.5  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931634.1  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931636.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932125.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932126.4  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932130.2

PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932186.8  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932206.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932227.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932228.7  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932229.5  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932230.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932922.2  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932926.5  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932928.1  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933004.2  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933005.0  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933006.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19940764.9  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940765.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940766.5  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940832.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19941378.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941379.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941380.0  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941394.0  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941396.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19942076.9  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942077.7  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942079.3  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942086.6  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942087.4  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942088.2  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942095.5  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942124.2  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942129.3  
PRIOR FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 1158  
SEQ ID NO 219  
LENGTH: 1557  
TYPE: DNA  
ORGANISM: Corynebacterium glutamicum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (101)..(1534)  
OTHER INFORMATION: RXN00132  
US-09-606-740A-219

Query Match 40.8%; Score 596.4; DB 23; Length 1557;  
 Best Local Similarity 64.5%; Pred. No. 4.6e-155;  
 Matches 923; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 32 tcgagtcgaatctgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 91  
 DB 114 tggacttcgaatctgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 173  
 QY 92 ctgagaagaatctgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 151  
 DB 174 cagagatgatctgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 233  
 QY 152 tgaaggtgtcaaatctgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 211  
 DB 234 tgaaggtgtcaaatctgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 293  
 QY 212 agacatcaagctcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 271  
 DB 294 agacatcaagctcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 353  
 QY 272 aagatcaagctcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 331  
 DB 354 agatcaagctcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 410  
 QY 332 gtaacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 391  
 DB 411 gtaacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 470  
 QY 392 gtccttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 451  
 DB 471 agatcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 527  
 QY 452 ctacatcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 511  
 DB 528 ccacatcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 587  
 QY 512 cagaagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 571  
 DB 588 aggcacatcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 647  
 QY 572 aagacacacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 631  
 DB 648 cagagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 707  
 QY 632 caacagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 691  
 DB 708 ccacagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 767  
 QY 692 tcaagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 751  
 DB 768 tgaagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 827  
 QY 752 cccttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 811  
 DB 828 cccttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 887  
 QY 812 tcaatgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 871  
 DB 888 tctgtgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 947  
 QY 872 ggtgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 931  
 DB 948 ggtgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1007  
 QY 932 agtcgcgcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 991  
 DB 1008 ctgtgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1067  
 QY 992 actgtgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1051  
 DB 1068 acaagacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1127  
 QY 1052 acatgtgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1111

DB 1128 acatgtgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1187  
 QY 1112 agcatcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1171  
 DB 1188 ccgcacacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1247  
 QY 1172 tcccttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1231  
 DB 1248 tcccttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1307  
 QY 1232 ttaagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1291  
 DB 1308 tcaagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1367  
 QY 1292 gaacgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1351  
 DB 1368 gacagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1427  
 QY 1352 tccacgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1411  
 DB 1428 tccacgttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1487  
 QY 1412 tcaagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1461  
 DB 1488 tccagcttcacatcttcgcgcacatcaactcctcgtcttcgcgcgaaggaacttacccttg 1537

RESULT 3  
 US-09-746-660A-97  
 ; Sequence 97, Application US/09746660A

; GENERAL INFORMATION:  
 ; APPLICANT: Pompeius, Markus  
 ; APPLICANT: Krogger, Burkhard  
 ; APPLICANT: Schröder, Hartwig  
 ; APPLICANT: Zelder, Oskar  
 ; APPLICANT: Habernauer, Gregor  
 ; APPLICANT: Kim, Jun-Won  
 ; APPLICANT: Lee, Heung-Schick  
 ; APPLICANT: Hwang, Byung-Joon  
 ; TITLE OF INVENTION: CORVNEBACTERIUM GLUTAMICUM GENES ENCODING  
 ; FILE OF INVENTION: METABOLIC PATHWAY PROTEINS  
 ; FILE REFERENCE: B61-121CP2

; CURRENT APPLICATION NUMBER: US/09/746, 660A  
 ; CURRENT FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 09/606740  
 ; PRIOR FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: 09/603124  
 ; PRIOR FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: 60/141031  
 ; PRIOR FILING DATE: 1999-06-25  
 ; PRIOR APPLICATION NUMBER: 60/142101  
 ; PRIOR FILING DATE: 1999-07-02  
 ; PRIOR APPLICATION NUMBER: 60/148613  
 ; PRIOR FILING DATE: 1999-08-12  
 ; PRIOR APPLICATION NUMBER: 60/187970  
 ; PRIOR FILING DATE: 2000-03-09  
 ; PRIOR APPLICATION NUMBER: DE 19931420.9  
 ; PRIOR FILING DATE: 1999-07-08  
 ; NUMBER OF SEQ ID NOS: 125  
 ; SOFTWARE: PatentIn Vers. 2.0  
 ; SEQ ID NO 97  
 ; LENGTH: 1557  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (101)..(1534)  
 ; OTHER INFORMATION: RXN01132  
 ; US-09-746-660A-97

Query Match 40.8%; Score 596.4; DB 29; Length 1557;

Best Local Similarity 64.5%; Pred. No. 4.6e-155;  
Matches 923; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 32 tcgagtagaattgcgcagatcaactcattcttcggtcgtgaaggtaactaccctg 91  
Db 114 tggactcaaggttcgcgactcttcactagcagggcgcgcgcacagatctgtctg 173  
QY 92 ctgagaagaataatgcaggtcttattgtcttcgtgagcggttattccgtcttaagcat 151  
Db 174 cagaatattagatgcaggtctcactagcttgcgaaggaattccgacagacgacct 233  
QY 152 tgaagggttcgaagatctcgttcctccacatgacagatccacagccgtccacacg 211  
Db 234 tgaagggttcgagatctcgttcctccacatgacagatccacagccgtccacacg 293  
QY 212 agacactcaagctctgtgtctgattcagatgaggttccttcgaacatctctctaac 271  
Db 294 agacctcaagctctgtgtctgattcagatgaggttccttcgaacatctctctaac 353  
QY 272 aagatagcgcgtctgtctgtctgtctgtccacagacacacagagaagccagcg 331  
Db 334 aggatgaggtctgcaggtctgtctgtctgtccacagacacacagagaagccagcg 410  
QY 332 gtatcccaagctctgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 391  
Db 411 gtgtcccaagctctgt 470  
QY 392 ggcgtctcaatggtccagatggttcaaggtccacagcaggtgtgtgtgtgtgtgt 451  
Db 471 agatctcagctgt 527  
QY 452 ctacactctcactccacaggttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 511  
Db 528 ccacacaggt 587  
QY 512 cagaagctgcagacactcgaatccagctgtgtgtgtgtgtgtgtgtgtgtgtgt 571  
Db 588 aggcacacagatccagatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 647  
QY 572 aagagaagaacacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 631  
Db 648 cagagctgcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 707  
QY 632 caacaggtgtcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgtgt 691  
Db 708 ccacacaggt 767  
QY 692 tcaacgtcagacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 751  
Db 768 tgaacgtcagacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 827  
QY 752 ccctatcgt 811  
Db 828 ccctatcgt 887  
QY 812 tcatggtgtcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgtgt 871  
Db 888 tctgtgtgtcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgtgt 947  
QY 872 ggcgttatacatcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgt 931  
Db 948 ggcgttatacatcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgt 1007  
QY 932 aggtcgc 991  
Db 1008 ctgtgtgtcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgtgt 1067  
QY 992 actgtgtgtcagacactgtgcacacaggtgtgtgtgtgtgtgtgtgtgtgtgt 1051  
Db 1068 acaagagcatcattctcctcgcagcagatgtgtgtgtgtgtgtgtgtgtgtgt 1127  
QY 1052 acatgcgcacactcgtatcagacacatgtgtgtgtgtgtgtgtgtgtgtgtgt 1111

Db 1128 acatgcgcacactcgtatcagacacatgtgtgtgtgtgtgtgtgtgtgtgt 1187  
QY 1112 agcacatcccatcagacacacacacacacacacacacacacacacacacacac 1171  
Db 1188 ccgcac 1247  
QY 1172 tccctctgtcagac 1231  
Db 1248 tgcgtcgtcagac 1307  
QY 1232 ttatgtatgtcattcagacacacacacacacacacacacacacacacacac 1291  
Db 1308 tcatgtcattcgtcagacacacacacacacacacacacacacacacacacac 1367  
QY 1292 gaaatcctcagac 1351  
Db 1368 gacagtcagac 1427  
QY 1352 tccacctcagac 1411  
Db 1428 tccacgtcagac 1487  
QY 1412 tcaacgtcagac 1461  
Db 1488 tgcgtcgtcagac 1537

RESULT 4

US-09-738-626-836  
Sequence 836, Application US/09738626  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIKI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738, 626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 836  
LENGTH: 1422  
TYPE: DNA  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-836

Query Match 40.7%; Score 594; DB 29; Length 1422;  
Best Local Similarity 64.5%; Pred. No. 2.1e-154;  
Matches 920; Conservative 0; Mismatches 500; Indels 6; Gaps 2;  
QY 32 tcgagtagaattgcgcagatcaactcattcttcggtcgtgaaggtaactaccctg 91  
Db 2 tggactcaaggttcgcgactcttcactagcagggcgcgcgcacagatctgtctg 61  
QY 92 ctgagaagaataatgcaggtcttattgtcttcgtgagcggttattccgtcttaagcat 151  
Db 62 cagaatattagatgcaggtctcactagcttgcgaaggaattccgacagacgacct 121  
QY 152 tgaagggttcgaagatctcgttcctccacatgacagatccacagccgtccacacg 211



















QY 962 atgctgatcttctgctacatgacagaagactgagatatactctgttgaatgttg 1021  
Db 15138 GCGGGCATCTATGTGACACACCGCAACAGACATCATACCGTGAGACCTGC 15259  
QY 1022 ccagatgaagataagactatgttgcgaacatcgccactcgaataatgata 1081  
Db 15258 AGGGATGAAAGACAGGCGATCTGTGCAACATCGGCCACTTCGACACAGATTCAGG 15199  
QY 1082 cagatgacctcaataatcccaagcgaataagacatcccaatcaagccagaataagaca 1141  
Db 15198 TCGATGCCCTGAATCACTTAAGGGGTGAGAAATCAACATCAACCCGAGTGCACA 15139  
QY 1142 tgtggaattccagatgagccagctatcctctcttctgtgagggcccttcaac 1201  
Db 15138 AGTATGTGTGGCAACGGTAACGCATCTCTCTGTGGCGAGCGCGCTGTGAACC 15079  
QY 1202 ttgctgctgaagtaagcccatcttctgtatgtcaatgttaataaaccagacac 1261  
Db 15078 TGGCGTGGCCACCGGCCACCGAGCTGTGTGATGTCCAACTGCTTGCACCAACAGACCC 15019  
QY 1262 tcgctcagctgagctctctacgaagaaggaatctcgaagaaggtttacacacttc 1321  
Db 15018 TGGCCAGATGCACTGTGGAAAGCGACAGCTACGAGAAAGAGTCTACATCTCTGC 14959  
QY 1322 cgaagcatcgtatgaagaagctgctgcctccactcgaatctcgtatgtccactta 1381  
Db 14958 CCACACCTGTGATGAAGAATGGCCCGCTGCACCTGGAAGATGGCGCTCAAGCTGA 14899  
QY 1382 caaagcttaacagaagaagctgactacatcaagctcagttgaggttcttaacagt 1441  
Db 14898 CCACCTGCAACGAAGCGCGGCGGCTGACCTGCGGAGAGCTGGCCGTAACAAGC 14839  
QY 1442 ctgatgtaccgttaata 1461  
Db 14838 CGGATCATTAACCGCTACTAA 14819

RESULT 14  
US-09-606-740A-223  
Sequence 223. Application US/09606740A  
GENERAL INFORMATION:  
APPLICANT: Pompejus, Markus  
APPLICANT: Krieger, Burkhard  
APPLICANT: Schroder, Hartwig  
APPLICANT: Zelder, Oskar  
APPLICANT: Habberhauer, Gregor  
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
METABOLIC PATHWAY PROTEINS  
FILE REFERENCE: BGI-121CP  
CURRENT APPLICATION NUMBER: US/09/606,740A  
CURRENT FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: 60/141,031  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 60/142,101  
PRIOR FILING DATE: 1999-07-02  
PRIOR APPLICATION NUMBER: 60/148,613  
PRIOR FILING DATE: 1999-08-12  
PRIOR APPLICATION NUMBER: 60/187,970  
PRIOR FILING DATE: 2000-03-09  
PRIOR APPLICATION NUMBER: DE 19930476.9  
PRIOR FILING DATE: 1999-07-01  
PRIOR APPLICATION NUMBER: DE 19931415.2  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931418.7  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931419.5  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931420.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931424.1  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931428.4  
PRIOR FILING DATE: 1999-07-08

PRIOR APPLICATION NUMBER: DE 19931434.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931435.7  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931443.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931453.5  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931457.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931465.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931478.0  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931510.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931541.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931573.6  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931592.2  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931632.5  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931634.1  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931636.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932125.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932126.4  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932130.2  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932186.8  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932206.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932227.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932228.7  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932229.5  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932230.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932922.2  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932926.5  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932928.1  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933004.2  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933005.0  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933006.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19940764.9  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940765.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940766.5  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940832.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19941378.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941379.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941380.0  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941394.0



PRIOR FILING DATE: 2000-06-23  
 PRIOR APPLICATION NUMBER: 60/141031  
 PRIOR FILING DATE: 1999-06-25  
 PRIOR APPLICATION NUMBER: 60/142101  
 PRIOR FILING DATE: 1999-07-02  
 PRIOR APPLICATION NUMBER: 60/148613  
 PRIOR FILING DATE: 1999-08-12  
 PRIOR APPLICATION NUMBER: 60/187970  
 PRIOR FILING DATE: 2000-03-09  
 PRIOR APPLICATION NUMBER: DE 19931420.9  
 PRIOR FILING DATE: 1999-07-08  
 NUMBER OF SEQ ID NOS: 125  
 SOFTWARE: PatentIn Vers. 2.0  
 SEQ ID NO 101  
 LENGTH: 1396  
 TYPE: DNA  
 ORGANISM: Corynebacterium glutamicum  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (101)..(1396)  
 OTHER INFORMATION: FRXA01371  
 US-09-746-660A-101

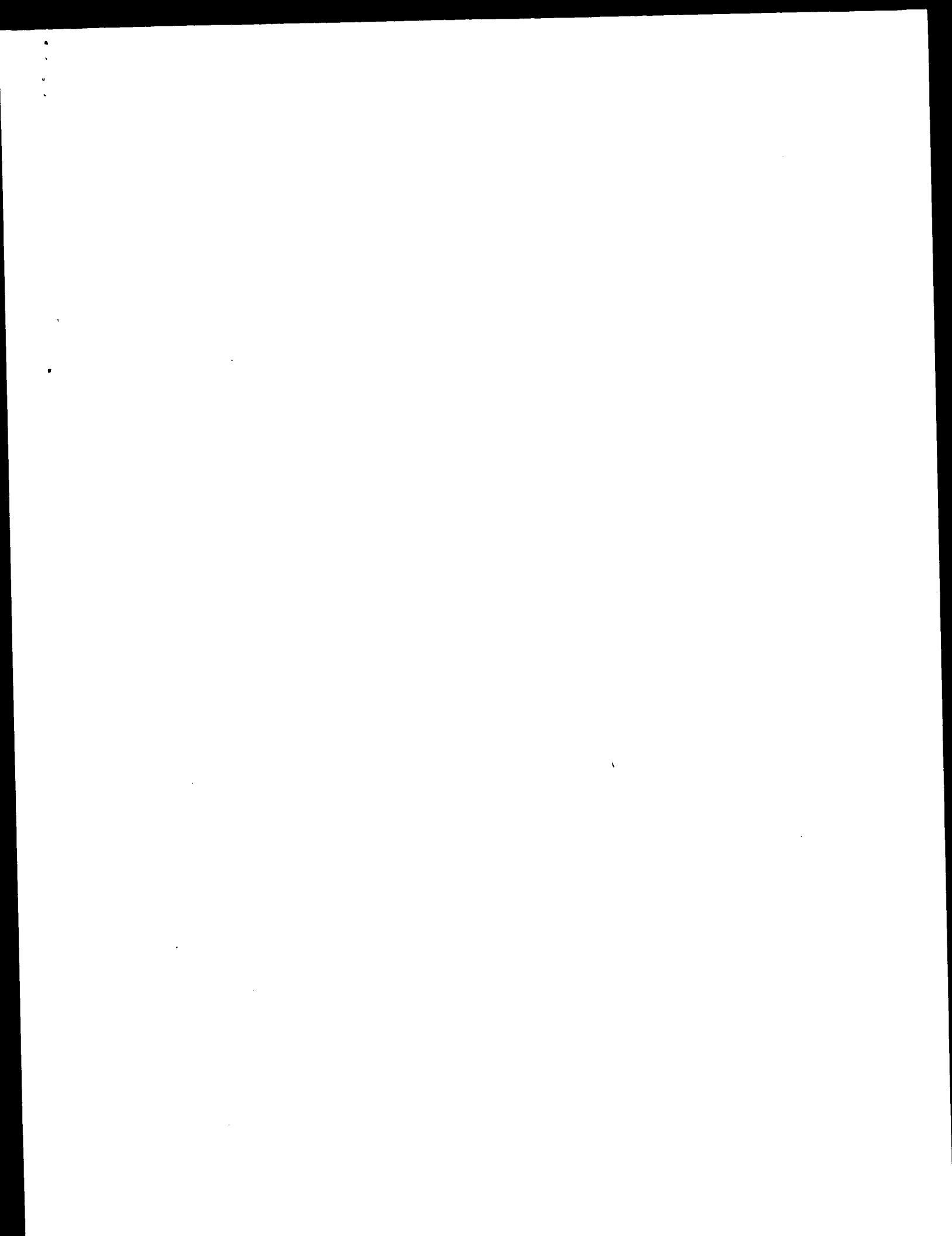
Query Match 36.2% Score 529.6; DB 29; Length 1396;  
 Best Local Similarity 64.3%; Pred. No. 1.9e-136;  
 Matches 828; Conservative 0; Mismatches 454; Indels 6; Gaps 2;  
 QY 32 tcgagtaacgaattgcgacatcaacccatcattctcgcgtaaggaactacccttg 91  
 DB 114 tggactcaaggttcgctcttcacatagcagagcagacgctcacaagattcttg 173  
 QY 92 ctggaagaagaatgcgacgtcttattgtcttcgtgagcgttactcgtcttcaagcat 151  
 DB 174 cagagtaagagatgcgacgtcttcaatgcagtcgcaagaatcgcgagagagagcctt 233  
 QY 152 tgaaggtgcgaatctctgtgtccctccacatgaacagtcgacagacgcgctctcatcg 211  
 DB 234 tgaagggcgccgaattgtgtgtcttctacacatgacgcgctcagacgcgctgtctatg 293  
 QY 212 agacactcaagctctgt 271  
 DB 294 agacactcaatctgt 353  
 QY 272 aagatcacgcgcgt 331  
 DB 354 agagtgagagctgcgacggt 410  
 QY 332 gtatccagctctcgt 391  
 DB 411 ggttcacagatctcgt 470  
 QY 392 ggcgtctcacatggtgcagatggtgcagagccacagcaggtgtgtgtgtgtgtgtgtgt 451  
 DB 471 agatcttcagctgt 527  
 QY 452 ctacactctctatctcgaaggt 511  
 DB 528 ccacatgt 587  
 QY 512 cagaagctacaaacctcgaatacgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 571  
 DB 588 aggcacaaatctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 647  
 QY 572 aagacaagaacacgtgcgacacagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 631  
 DB 648 cagagctgt 707  
 QY 632 caacaggt 691  
 DB 708 ccacgcgt 767  
 QY 692 tcaagtcacagcgcgt 751

DB 768 tgaagtcacagcgcgt 827  
 QY 752 cccattacagatgtatcaaacggt 811  
 DB 828 cccattacagatgtatcaaacggt 887  
 QY 812 tcatgt 871  
 DB 888 tctgt 947  
 QY 872 ggtttatcatcaacagactcgaacatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 931  
 DB 948 ggttcacagcgcgt 1007  
 QY 932 aggtccgcgcgt 991  
 DB 1008 ctgt 1067  
 QY 992 actgcgatacatctctgt 1051  
 DB 1068 acaagacatcatctctgt 1127  
 QY 1052 acatgcgcacatctgt 1111  
 DB 1128 acatgcgcacatctgt 1187  
 QY 1112 agcacatcccaatcaacaggaatcagacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1171  
 DB 1188 ccgcacacacatcaacaggaatcagacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1247  
 QY 1172 tctctgt 1231  
 DB 1248 tctgt 1307  
 QY 1232 ttatgtcatgt 1291  
 DB 1308 tcatgt 1367  
 QY 1292 gaaatctcgaagaagaagtttacacact 1319  
 DB 1368 gacagtcagagaagaggtctacacgtct 1395

Search completed: August 17, 2002, 04:32:22  
 Job time: 9944 sec

Mon Aug 19 15:16:25 2002

us-09-759-990-1.rnp







9 AspPhelysValAlaAspleuSerleuAlaIaphGlyArgLysGlu11 25  
 84 TACCTTGGCTGAGAGAAATGCCAGCTTATGGTTCTTGAGGCGT 133  
 25 ePrlleuAlaGlnHisGluMetProGlyLeuMetAlaIleArgLysGluT 42  
 134 ATTCGGCTTAAAGCCATGTGAAGGCTGCATGACCTCTGGTCCCTCCAC 183  
 42 YrAlaGlnAlaGlnProleuAlaGlnAlaArgValThrGlySerleuHis 58  
 184 ATGACAGTCCAGACAGCGCTCCATCGAGACACTCACAGCTCTTGCTGC 233  
 59 MetThrValGlnThrAlaValLeuIleGlnThrLeuValAlaLeuGlyAl 75  
 234 TGATGTCAGATGGCTCTCCATCATCTCTTACACAGATACAGCGCG 283  
 75 AgluValArgTrrAlaSerCysAsnIlePheSerThrGlnAspHisAla 92  
 284 CTCTGTATCGTGTGCGCCCAACAGGACACAGACAGAGAGCCGCGT 333  
 92 IAlaIAlaIleAlaValGlyProAsnGlyThrProAspAsnProGlnGly 108  
 334 ATCCCACTCTTCCCTGGAAGGCGAAACACTCCCAATACTGGGAGAA 383  
 109 ValProValIleAlaTrpLysGlyLthrLeuGluGlyTrpTrpCys 125  
 384 CAGATACCGGCTCCATCGACGACGATGTCA.....GGCCACAGC 427  
 125 sTrpGluGlnAlaLeuThrTrpProAsnThrProThrGlyGlyProAsn 142  
 428 AGGTGTGATGATGTGGTATGCTACACTCTTCATCTCCAGGCGTTC 477  
 142 eTlleuAspAspGlyGlyAspAlaThrLeuValHisLysGlyVal 158  
 478 GAATTCGAACAGCGCGCTGCTCCAGACGCAACAGAAAGTCGACAGCT 527  
 159 GluTrpLysAspGlyLysValProSerValAspThrAlaGlnSerHis 175  
 528 CGAATACCGGCTCTGCTTACACTCAAGAGGCTTCAACCAAGACA 577  
 175 pGlnHisArgValIleLeuGluLeuLeuThrArgThrValGlyGluSer 192  
 578 AGAACCACTGGACACAGTGTCTGCGCATGACAGGCTTCCGAGAG 627  
 192 roGlnLysTrpThrGlnLeuAlaSerGluLeuArgLysValThrGluGln 208  
 628 ACAACACAGGCTGTCACCGCTTACAGCTCGAAGAGGAGGCAACT 677  
 209 ThrThrThrGlyValHisArgLeuTrpGluMetHisArgAspLysThr 225  
 678 CTTCTCCACAGCCATCAACGTCACAGCGCTGTACAAGTCCAACTGCG 727  
 225 uIleuPheProAlaIleAsnValAsnAspAlaValThrLysSerLysPhe 242  
 728 ATAACATCTACGCTGCGCGCACTCCCTTATGATGGTATCAACCTGCT 777  
 242 spAsnLysTrpGlyCysArgHisSerLeuIleAspGlyLleAsnArgAla 258  
 778 TCCCATCTCATGCGCGCGCAAGACACTCTCGATGGTATCCGCGCA 827  
 259 ThrspAlaLeuIleGlyGlyLysThrAlaValAlaCysGlyLysLys 275  
 828 TGTGCGCAAGGCTGCGCTCAATCCCTCGGCGCAAGGCGCTGCGCTTA 877  
 275 pValGlyLysGlyCysAlaGlnSerLeuArgGlyGlnGlyAlaArgVal 292  
 878 TCATCACAGAACTGACCAATCTGCGCTCCAGCGCTCCATGGAAGC 927  
 292 IeIleThrGlnIleAspProIleCysAlaLeuGlnAlaIleMetAspGly 308  
 928 TACCAAGTCCGCGCATCGAGAACTGTCGAAGATGTCGATATCTTCGT 977  
 309 PheGlnValThrThrLeuAspGluValAlaAspLysAlaAspIlePhe 325

978 TACATGCAGAAACTGCATATCATCTGTGTGACATGATGCCAGCA 1027  
 325 LThrThrThrLysAsnLysAspIleIleMetAlaLysAspMetAlaLys 342  
 1028 TGAAGATGAGCTATTGTGCGTAAACATCGCCACTTCATTAAGAAAT 1077  
 342 eLysHisGlnAlaIleValGlyAsnIleGlyHisPheAspAsnGluIle 358  
 1078 GATACAGATGGCTTCATGAATACCCAGGATCAACACATCCCAATCA 1127  
 359 AspMetAlaGlyLeuAlaGlnThrProGlyIleValLysAspLysVal 375  
 1128 GCCAGATACGACATGTGGGAATTCACAGATGCCAGCACTATCTCTTC 1177  
 375 sProGlnValHisThrTrpThrTrpProAspGlyLysValLeuIleVal 392  
 1178 TTGCTGAGGCGCGCTTCTTAACCTTGGCTGGCTACAGTCAACCATCT 1227  
 392 euserGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 408  
 1228 TTGCTTATGTCATGTCATTCACAAACAGACACTGCTCAGCTGACACT 1277  
 409 PheValMetSerAsnSerPheAlaAspGlnThrLeuAlaGlnIleGlu 425  
 1278 CTACGAAAGAGAGAAATCTCGAAGAGGTTTACACACTTCGAGAC 1327  
 425 uPheThrLysProAspArgLysTrpProThrAspValTyrValLeuProLys 442  
 1328 ATCTCGATGAAGAACTGCTGCTCCAGCTCGATCTCGATGTCAC 1377  
 442 ILeuAspGluLysValAlaArgLeuHisLysAspLeuGlyValLys 458  
 1378 CTTACAAAGCTTACACAGAGCGCTGACTACATCACTTCACATTTGA 1427  
 459 LeuThrThrLeuArgProGluGlnAlaAspTyrIleGlyValLysVal 475  
 1428 GGGCTCTTACAAAGTGTGATGCTTACCGTTAT 1458  
 475 uGlyProTyrLysAlaAspHisLysTyrArgTyr 485

seq\_name: sp\_bacteria:Q936D6

seq\_documentation\_block: PRT: 482 AA.  
 ID Q936D6; PRELIMINARY;  
 AC Q936D6;  
 DT 01-DEC-2001 (TREMBLER, 19, Created)  
 DT 01-DEC-2001 (TREMBLER, 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)  
 DE ADHICASE.  
 GN MTM.  
 OS Streptomyces argillaceus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=41951;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fernandez Lozano M., Lombo F., Brana A.F., Sales J.A., Mendez C.;  
 RT "Generation of a Streptomyces argillaceus mutant by deleting two  
 RT methyl cycle-related genes located within the mlbA gene cluster  
 RT leads to higher production of methylamylol."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ16377; CAC94890.1; B3C4D7ADB25606 CRC64;  
 SQ SQUENCE 482 AA; 52563 MW; B3C4D7ADB25606 CRC64;

alignment\_scores:

Quality: 1625.50 Length: 477  
 Ratio: 4.064 Gaps: 2  
 Percent Similarity: 83.857 Percent Identity: 65.828

alignment\_block:

us-09-759-990-1 x Q936D6 ..

Align seg 1/1 to: Q936D6 from: 1 to: 482

```

34 GAGTACAGAAATGGCCGACATCAACATCCATGTTCTGGCCGCTAAGAACT 83
   |||||
7  AspPhelysValAlaAspIleuSerLeuAlaGluPheGlyArgGlyGluI 23
84 TACCTTCGTCGACAGAAATGCCAGGCTTATGGTCTTCTGAGCGCT 133
   |||||
23 eThrLeuAlaGluHisGluMetProGlyLeuMetSerIleArgArgGluT 40
134 ATTCGCGTTTAAAGCCATTTGAAGGTTGTCAGATCTCGTCCCTCCGAC 183
   |||||
40 yTrAlaGluAlaGluProLeuAlaGlyAlaArgValThrGlySerLeuHis 56
184 ATGACAGTCCAGACAGCCGCTCCATGAGACATCCAGACAGCTTGTGTGC 233
   |||||
57 MetThrValGlnThrAlaValLeuIleGluThrLeuValAlaLeuGlyAl 73
234 TGGATGTCAGATGGGCTTCCTGCAACATCTTCTTACACAGATACAGCCG 283
   |||||
73 aGlnValAlaGlyTrpAlaSerCysAsnIlePheSerThrGlnHisAla 90
284 CTGCTGCTATGCTGTTCGCGCCACAGAGGACACAGAGGACAGCCGCT 333
   |||||
90 IalaIalaIleAlaValGlyProAspGlyThrProAspAsnProArgAla 106
334 ATCCAGCTCTTCGCTGGAAAGGCGAAACATCCAGAAATACGGAGAA 383
   |||||
107 ValProValPheAlaTrpLysGlyGluThrLeuGluGluGlyTrpTrp 123
384 CACATACCCCGCTTCACATGCCAGATGGCA.....GCCGACAGC 427
   |||||
123 sThrGluGlnAlaLeuThrTrpProAspSerProThrGlyIleProAsn 140
428 AGCTGTGTCATGATGGTGTGATGCTACATCTCCATCTCCAAAGGCGCTC 477
   |||||
140 eTrIleuAspAspGlyArgLysAlaThrIleLeuValHisLysGlyAla 156
478 GAATTCGAAACAGCCGCTGCTTCAGAGCCAAAGAAAGTGCACACCT 527
   |||||
157 GluTrpGluLysAspGlyLysValProSerValAspThrAlaGluAsnAs 173
528 CGAATTCGCGCTGCTTCTACATCAAGAGGCTTCAACACCAAGACA 577
   |||||
173 pGluHisAlaGlyValIleLeuGluLeuAsnAspArgThrIleThrAsp 190
578 AGAACCACTGGCACACAGTGTGCGCGCATGAAGGCTGTTCGGAAGAG 627
   |||||
190 eTrGlnLysTrpThrGlnLeuAlaSerGluIleArgGlyValThrGluGlu 206
628 ACAACAAAGGCTGTCACCGCTCTTACAGCTCGAAGAGGCGCAACT 677
   |||||
207 ThrTrpThrGlyValHisArgLeuTrpGluMetIleArgAspGlyValLe 223
678 CTTTCTCCAGCATCAAGCTCAAGCGGCTTAAAGAGCCAGTCCG 727
   |||||
223 uLeuPheProAlaIleAsnValAsnAspAlaValThrLysSerLysPheA 240
728 ATACATCATAGCGCTGCCGACATCCCTTATCATGATGATCAACCGCT 777
   |||||
240 sPasnLysTrpGlyLysArgHisSerLeuIleAspGlyIleAsnArgAla 256
778 TCCGATGTCATATGCGCGCAAGACATCTGTATGGGTTACGGCGA 827
   |||||
257 ThrAspValLeuIleGlyLysThrAlaValAlaPheGlyTrpGlyAs 273
828 TGTGCGAGGCGCTGCGTCATACCTCCGCTGGCGAAGGCGCTGTA 877
   |||||
273 pValGlyLys...AlaIleArgSerLeuArgGlyGlnGlyAlaTrpValI 289
878 TCATATCAGAACTGCACCAATCTGCGCTTCCAGGCTGCCATGGAAGC 927
   |||||

```

```

289 IeValThrGluIleAspProIleCysAlaLeuGlnAlaIleMetAspGly 305
928 TACGAGTCCGCGCATGCGAAGTCGTCAAAGATGTCATATCTTCGT 977
   |||||
306 TyrGlnValThrThrLeuAspGluValAlaAspLysAlaSerIlePheI 322
978 TACATGCACAGAAACTGCCATATCATCTCTGTGACATGATGGCCAGA 1027
   |||||
322 eThrThrThrGlyAsnLysAspIleIleMetAlaIleAspMetAlaArgM 339
1028 TGAAGATTAAGCTTATGTCGATACATGCGCCATTCGATTAAGAAATT 1077
   |||||
339 eLysHisGlnAlaIleValGlyAsnIleGlyHisPheAspAsnGluIle 355
1078 GATACAGATGGCTCATGAAATATCCAGAGCATCAAGACATCCCAATCA 1127
   |||||
356 AspMetAlaGlyLeuAlaLysValProGlyIleValLysAspGluIle 372
1128 GCCAGAAATACGACATGCGGAATCCAGATGGCCAGCAGCTATCCCTTC 1177
   |||||
372 sProGlnValHisThrTrpThrPheProAspGlyLysIleIleVal 389
1178 TTGCTGAGGCGCGCTTCTTAACTTGGCTGCGCTACAGTCAACCATCT 1227
   |||||
389 eUserGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlnHisProSer 405
1228 TTGCTTATGATGTCATTCACAAACAGACATCCGCTCAGCTGACACT 1277
   |||||
406 PheValMetSerAsnSerPheAlaAspGlnThrLeuAlaIleGluLe 422
1278 CTACGAAAGAGAGAAATCTCGAGAAAGGTTTACACATCTCCGAAAC 1327
   |||||
422 uPheThrLysProGlyAspTrpProThrGlyValGlyThrLeuProLysH 439
1328 ATCTGATGAGAAAGTCCGCTCCGCTCCACCTCGATCTGCTGTCAC 1377
   |||||
439 IsteAspGluLysValAlaArgLeuHisLeuAspAlaLeuGluLys 455
1378 CTTACAAAGCTTACACAGAGCGCTGACATCATGATGATGCTGATGTA 1427
   |||||
456 LeuThrThrLeuArgProGluGlnAlaSerTrpIleGlyValAspValAs 472
1428 GCGTCTTACAAAGTGTGATGCTTACCGTTAT 1458
   |||||
472 pGlyProGlyLysSerAspHisLysTrpArgTrp 482
seq_name: sp_bacteriap:Q9CCJ4
seq_documentation_block:
ID Q9CCJ4 PRELIMINARY; PRY; 492 AA.
AC Q9CCJ4;
DT 01-JUN-2001 (TREMblrel. 17, created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE PUTATIVE S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE.
GN SAH OR M0771.
OS Mycobacterium leprae.
OC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
OC Actinomycetales: Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Dutfoy S., Feltham T., Fraser A., Hamlin N.,
RA Hotroyd S., Hornsby T., Jagsels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Squires S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.,

```

"Massive gene decay in the leprosy bacillus."  
 RT Nature 409:1007-1011(2001).  
 RL EMBL: A1583919; CAC30280.1;  
 DR HSSP; P10760; 1D4F.  
 DR Leproma; ML0771;  
 DR InterPro; IPR000043; Ado\_hcyase.  
 DR Pfam; PF00670; AdoHcyase; 1.  
 DR PROSITE; PS00738; ADHCHYASE\_1; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 492 AA; 53986 MW; DE83D454CF87EB5 CRC64;

alignment\_scores:  
 Quality: 1537.50 Length: 477  
 Ratio: 3.932 Gaps: 3  
 Percent Similarity: 81.971 Percent Identity: 62.264

alignment\_block:  
 US-09-759-990-1 x Q9CCJ4 ..

Align seg 1/1 to: Q9CCJ4 from: 1 to: 492

```

34 GAGTACAGAAATTCGGACATCACTCCATCTTCGCGCCCTAAGAACT 83
   ::::::::::::::::::::
17 AspHelysValAlaAspLeuSerLeuAlaAspHelysValAlaGlySglu 33
   TACCCCTGTCGAGAAATCCAGCTCTTATGCTTCGTCGACGCTT 133
   ||||||| ||||||| ||||||| ||||||| |||||||
33 uAspLeuAlaGlyTyrGluMetProGlyLeuMetSerLeuAlaGlySglu 50
   ||||||| ::::::::::::::::::::
134 ATTCGCTTGAAGCATTGAAGGTGTCAAAATCTGCTGCTCCAC 183
   ::::::::::::::::::::
50 yAlaGlyValAlaGlnProLeuSgluAlaArgIleSerGlySerLeuHis 66
184 ATGACAGTCCAGACGCGCTCCATCCAGACATCCAGACATCCAGCTTGGTG 233
   ||||||| ||||||| ||||||| ||||||| |||||||
67 MetThrValGlnThrAlaValAlaLeuIleGlnThrLeuThrAlaLeuSglu 83
   ||||||| ||||||| ||||||| ||||||| |||||||
234 TGATGTCAAGATGGCTCTCCAGACATCTTCTTACACAAATACAGCCG 283
   ||||||| ||||||| ||||||| ||||||| |||||||
83 aGlyValAlaArgTyrAlaSerCysAsnIlePheSerThrGlnAspHisAla 100
284 CTGCTGCTATCTGTGTGCGCCACACAGGACACACAGAAAGCCGCT 333
   ||||||| ||||||| ||||||| ||||||| |||||||
100 lAlaAlaValAlaValAlaGlyProTyrGlyThrProGlnGluProLysGly 116
334 ATCCAGATCTTGCCTGGAAGGCGCAACATCCACAAATACCTGAGAA 383
   ||||||| ||||||| ||||||| ||||||| |||||||
117 ValProValPheAlaTyrPheLysGlyThrLeuGlnGluTyrTrpThrAl 133
384 CACATACCGCGCTCTCACATGCGCCAGATGTCAAGGCCA...CAGCAGG 430
   ::::::::::::::::::::
133 aAlaGlnGlnMetLeuThrTrpProAspProAspLysProValAlaMetI 150
431 TTGTCGATGATGATGTCGATGTCATCACTCCATCCAGAGGCTCGAA 480
   ::::::::::::::::::::
150 lLeuAspAspGlyGlyAspAlaThrMetLeuValAlaLeuArgLysValGln 166
481 TTGCAACACAGCGCTGCTGCTCCAGACCAACAGAACTGACCAACCTGA 530
   ::::::::::::::::::::
167 TyrGlnLysAlaGlyValAlaProProAlaGlnValAlaAspSerAlaGln 183
531 ATACCGCTGCTGCTTGTCTTACACTCAAGCAGGCTTCAACCAAGACAGA 580
   ::::::::::::::::::::
183 uTrpLysValAlaPheLeuAsnLeuLeuArgLysArgPheGlnThrAspLys 200
581 ACCACTGACACAGTTCGTCGCGCATGAGGCTTTCGGAAGAGACA 630
   ::::::::::::::::::::
200 lLysTrpThrLysIleAlaLysSerValLysGlyValThrGlnGluThr 216
631 ACACAGAGTTCACCGCCTTACACAGCTCCAGACAGAGGAGCAAACTCT 680
   ||||||| ||||||| ||||||| ||||||| |||||||
217 ThrThrGlnValLeuArgLysLeuTyrGlnPheAlaAlaAlaGlyAspLeuAl 233

```

```

681 CTTCACAGCCATCAAGCTCAACAGAGCTGTACAAAGTCCAAAGTTGATCA 730
   ||||||| ||||||| ||||||| ||||||| |||||||
233 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 250
731 ACATTCACAGCTCCGCGCTCTTATCGATGGTATCAACCGGCTCC 780
   ||||||| ||||||| ||||||| ||||||| |||||||
250 snLysTyrGlyThrAlaGlnSerLeuIleAspGlyIleAsnAlaGlyThr 266
781 GATGTCATGATCGCGCAGACAGCTCTGTCATGGGTTCAGCGGATGT 830
   || ::::::::::::::::::::
267 AspSerLeuIleGlyGlyLysAsnValLeuIleCysGlyTyrGlyAspVa 283
831 CGGCAGAGGCTGCGCTCATCCCTCCGTCGCCAAGCGCTGCGTTATCA 880
   ||||||| ||||||| ||||||| ||||||| |||||||
283 lGlyLysGlyCysAlaGlnAlaAlaLysGlyGlnGlyAlaArgValThrI 300
881 TCACAGACCTCCAGCCATCGGCTCCAGGCTCCAGGCTCCAGGCTAC 930
   ||||||| ||||||| ||||||| ||||||| |||||||
300 lThrGlnIleAspProIleAsnAlaLeuGlnAlaLeuMetGlnLysPhe 316
931 CAGTCCCGCATCGAAGAGTGTCAAGGATGTGATATCTTCGTTAC 980
   ::::::::::::::::::::
317 AspValLysArgValAlaAspValIleAlaAspSerAlaValAlaThr 333
981 ATGCACAGAAATCGCATATCATCTGTTGACATGATGCCCATGA 1030
   ||||||| ||||||| ||||||| ||||||| |||||||
333 rAlaThrGlyAsnLysAspIleIleLeuIleGlnHisMetLysAlaMetL 350
1031 AGGATAGGCTATGTCGTAACATCGGCGCATTCGATTAAGAAATGAT 1080
   ||||||| ||||||| ||||||| ||||||| |||||||
350 yAspHisAlaIleLeuGlnLysnIleGlyHisPheAspAsnGluLeuAsp 366
1081 ACAGATGCGCTCATGAATCCAGAGCATCAACATCCAAATCAAGCC 1130
   ::::::::::::::::::::
367 MetAlaAlaLeuGlnArg...SerGlyAlaThrArgLeuAsnIleLysBr 382
1131 AGAATACGACATGTGGGAATCCAGAT...GCCAGCCTATCCCTCTTC 1177
   ||||||| ||||||| ||||||| ||||||| |||||||
382 OGlnValAspLeuThrPheGlyAspSerGlyLysSerIleIleVal 399
1178 TTGTCGAGGCGCTTCTTAACCTTGCTGCTGCTACAGTCCACCATCT 1227
   ||||||| ||||||| ||||||| ||||||| |||||||
399 euserGlnGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProser 415
1228 TTGCTATGTCATGTCATTCACAAACACACACTGCTCAGCTGACCT 1277
   ||||||| ||||||| ||||||| ||||||| |||||||
416 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGlu 432
1278 CTACGAAGAAGAGAGAAATTCGAGAGAGAGTTTACACACTTCGGAAC 1327
   ||||||| ||||||| ||||||| ||||||| |||||||
432 uTrpThrLysAsnAspAspTyrAspAsnGlnValAlaTyrArgLeuProLysH 449
1328 ATCTCGATGAAGATGCTGCTGCTCCACCTCGATGCTCGATGTCAC 1377
   ||||||| ||||||| ||||||| ||||||| |||||||
449 lLeuAspGlnLysValAlaArgValAlaHisValGlnAlaLeuGlyGln 465
1378 CTTCACAAAGCTTACACAGACAGGCTGACTCATCAACGCTCCAGTTCA 1427
   ||||||| ||||||| ||||||| ||||||| |||||||
466 LeuThrLysLeuThrLysAspGlnAlaGlnLysGlyValAlaAspVal 482
1428 GGGTCCTTACAAAGTGTGCTTACCGTTAT 1458
482 pGlyProPheLysProAspHisTyrArgTyr 492

```

seq\_name: sp\_plant:Q9LK36

seq\_documentation\_block:  
 ID Q9LK36 PRELIMINARY; PRT; 485 AA.  
 AC Q9LK36;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)





```

282 ytrhlaaglyalargvalililevalhrgluilileasprroillecyssala 298
907 TCAGAGGTCGTCATGAGGAGGTCACAGGTCGCCGCGCATGAGAGAGTCGT 956
299 LeuGlnAlaLeuMetGluGluGluGlnValLeuThrLeuGluAlaValVa 315
957 CAAGATGTGATATCTCTGTTACATGCACAGAAATGCGATTCATCT 1006
315 lsergualaasprilephecysrthrthrthrglyasnlyasprillelem 332
1007 CTGTGACATGATGCGCCAGATGAAGATGAAGATGATGTCGTATACATC 1056
332 etvalasprhismetarglysmetlyssasnasnalalevalcysasnile 348
1057 GCGCACTTGATGATGAGAAATGATATACAGATGAGTCATGAAATACCCAG 1106
349 GlnHisPheasprasnngluileasprmetleuGluThrThrProcl 365
1107 CATCAAGCAACATCCCAATCAAGCCAGAAATACGACATGCGGAATCCAG 1156
365 yalllyasargilethrilleysproglthrAspargtrrvalpneproa 382
1157 ATGGCCACGCT...ATCCCTCTCTGTGTGAGGCGGCTCTCTTAACCTT 1203
382 sPrThasnserrglylellevalleuAlaGluGlyargleuMetlsnleu 398
1204 GCGTGGCTACAGATGACCCATCTTCTGTTATGTCATGATCATTCACAAA 1253
399 GlyCysAlaThrGlnHisProserPheValMetserCysSerPheThrAs 415
1254 CGAGACATCTGCTGAGTGCAGCTCTTACAGAAAGAGA.....GGAATTC 1297
415 nglnvalillealaglnleuGlnleuThrPasnngluysSerSerGlyLysT 432
1298 TCGAGAGAAGGTTTACACACTTCGAGATCGAGATCGATGAGAGAGTCGT 1347
432 yrglulyslysaltyvalleuProlyshisLeuAspGluysvalAla 448
1348 GCGCTCCACCTCGATCTCTGATGATCCACCTTACAAAGCTTACACAGAA 1397
449 AlaLeuHisLeuGluGluysLeuGluAlaArgleuThrLysLeuThrLysAs 465
1398 CGAGGCTGACATACAGCTCCAGTTCAGAGGCTCTTACAAAGCTGATG 1447
465 pElnserrAsrtyValserlleProvalGluGlyProtyGlyLysProvalH 482
1448 CTTACCGTTAT 1458
482 lSTYATrGtyr 485
seq_name: sp_plant:Q9SP37
seq_documentation block:
ID Q9SP37 PRELIMINARY; PRT; 485 AA.
AC Q9SP37;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE
DE HYDROLASE) (ADHOCYASR).
GN SHH.
OS Lupinus luteus (yellow lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
OX NCBI_TaxID=3873;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOTS INFECTED WITH BRADYRHIZOBIUM LUPINI;
RA Janowski R., Podkowiński J., Kisiel A., Jaskolski M.;
RT "S-adenosyl-L-homocysteinease cDNA sequence from Lupinus luteus.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

```

```

CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE
CC + L-HOMOCYSTEINE.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
DR EMBL: AF185635; AAD56048.1; -.
DR HSSP: P10760; 1BRF.
DR InterPro: IPR000043; Ado_hcyase.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000860; Porphobil_deam.
DR Pfam: PF00670; AdhCysase.1.
DR PROSITE: PS00738; ADHOCYASE.1.1.
DR PROSITE: PS00739; ADHOCYASE_2.1.
KW Hydrolyase; NAD; One-carbon metabolism.
SQ SEQUENCE 485 AA; 53326 MW; 42F8PA658A42B19 CRC64;

```

## alignment\_scores:

```

Quality: 1506.00 Length: 482
Ratio: 3.862 Gaps: 4
Percent Similarity: 80.913 Percent Identity: 61.618

```

## alignment\_block:

```

US-09-759-990-1 x Q9SP37

```

```

Align seg 1/1 to: Q9SP37 from: 1 to: 485

```

```

34 GAGTACAGAAATGGCCGACATCAACCTCCATGCTTCTCGGCGTAAGAACT 83
13 GluThrLysValLysAspMetSerGlnAlaSprhGlyArgLeuLui 29
84 TACCTTGGCTGACAGAAATGCGAGTCTTATGCTTCTTCTGACGCTT 133
29 egulEuAlaGluValGluMetProGlyLeuMetAlaSerArgSerGluP 46
134 ATTCGCGTCTTAAAGGTCATGAAGGTCGACAGATCTGTCGTCGTCGTC 183
46 heGlyProSerGlnProPheLysGlyAlaLysIleThrGlySerLeuHis 62
184 ATGACAGTCCGACAGAGCGGCTCTCATCGAGACATCAGACGCTTGGTGC 233
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGluAl 79
234 TGATGTCAAGATGGGCTTCTGCAACATCTTCTTACACAGATACAGCCG 283
79 agluValArgTrpCysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATGCGTTGTGCGCCCAACAGGACACACAGAGAAAGCCAGCG 333
96 lAlaAlaIleAlaAlaArgAspSerAlaAla..... 105
334 ATCCAGATCTTCTGCTGGAAGGCGGAACACTCCAGATATAGTGGAGAA 383
106 .....ValrPheAlaTrpLysGluThrLeuGlnGluThrTrpTrpCys 120
384 CACATACCGCGCTCTCAATGCGACAGATGTCAGAGGCCACAGACAGCTTG 433
120 sThrGluArgAlaLeuAspTrpGlyProGlyGlyGlyProAspLeuIleV 137
434 TCGATGATGGTGGATGATGATCTCCATCTCCATCCAGAGGCTTGGAA... 480
137 alAspAspGlyGlyAspThrThrLeuLeuHisGluGlyValLysAla 153
481 .....TTGAAACAGCGGCTGCTTCTTCCAGACGACCAAGAACTGA 521
154 GluGluIleLysGlyLysSerGlnPheProAspProAspSerThrAs 170
522 CAACCTCGAATACCGGCTGCTTCTTCTTACACTCAAGCAGGTCCTTCAACC 571

```







13 GluTyrLysValLysAspMetSerGlnAlaAspPheGlyArgLeuGluI 29  
 84 TACCCCTGGTGAAGAAGAAATGCCAGGCTCTTATGTTCTTCTGAGCGTT 133  
 29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaSerArgAlaGluP 46  
 134 ATTCGGCTTCTAAGCCATTGAAAGGGTGTCAAGATCTGTGGTCCCTCCAC 183  
 46 heGlyProSerGlnProValLysGlyAlaLysIleThrCysSerLeuHis 62  
 184 ATGACAGTCCAGACAGCCGCTCCATGAGAGCACTCAAGCTCTTGATGTC 233  
 63 MetThrIleGlnThrAlaPheLeuIleGluThrLeuThrAlaLeuGlyAl 79  
 234 TGATGTCAAGATGGCTTCCCTGCAACATCTTCTTACACAGATACAGCCG 283  
 79 agIuValAlaGlyTrpCysSerCysAsnIlePheSerThrGlnAspHisAla 96  
 284 CTGCTGCTATCTGTGTGCGCCCAACAGAGCACACAGAGAACCCGCGGT 333  
 96 laIAlaIleAlaIleAlaArgAspSerAlaAla..... 105  
 334 ATCCAGTCTCTGCTGGAAGGCGCAAAACACTCCAGAAATAGTGGAGAA 383  
 106 .....ValPheAlaIleTrpLysGlyGluThrLeuGlnGluTyrTrpTrp 120  
 384 CACATACCGCGCTCTCAATGCGCCAGATGTCACAGGCCACAGCAGTTG 433  
 120 sThrGluArgAlaLeuAspTrpGlyGlyGluValGlyProAspLeuIleV 137  
 434 TCATGTATGGTGGTGTGATGCTACACTCTCATCTCCAAAGGGCTTC..... 477  
 137 alAspAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 153  
 478 .....GAATTCGAACAGCGGCTGTCTGTCCAGAGCCACAGAACTGA 521  
 154 GluGluGluPheGlnLeuValLeuThrIleIleLysGluSerLeuLysT 170  
 522 CAACCTGGAATACCGCTGCTTGTGCTTACACTCAAGCAGAGTCTTCAAC 571  
 170 pasValGluPheGlnLeuValLeuThrIleIleLysGluSerLeuLysT 187  
 572 AAGCAAGAACCACTGGCAACAGTGTCTGCGCGCATGAACGGTGTTC 621  
 187 hrAspProLeuArgTyrThrLysMetLysGluArgLeuValGlyValSer 203  
 622 GAAGAGACAAACAGGTGTCCACCGGCTCTTACAGCTCGAAGAGAGGG 671  
 204 GluGluThrThrThrGlyValLysLysLeuTyrGluMetProAlaAsnG 220  
 672 CAACCTGCTCTCCAGCCATCAAGCTCAAGCAGCTGTATCAAACTCA 721  
 220 ySerLeuLeuPheLeuProIleAsnValAsnAspSerValThrLysSerL 237  
 722 AGTTGATTAACATCTACGGCTGCGCCACTCCCTTATCATGTGATTAAC 771  
 237 yspheAspAsnLeuTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253  
 772 CGTGTTCGATGTCATGCGCGGCAAGACAGCTCTCGTCAATGGTTA 821  
 254 ArgAlaThrAspValMetIleAlaGlyLysValAlaLeuValAlaGlyT 270  
 822 CGGCGATGTGCGCAAGGGCTGCGCTCAATCCCTCGTGGCCAGCGCTC 871  
 270 rGlyAspValGlyLysGlyAlaAlaMetLysGlnAlaGlyAlaAla 287  
 872 GCGTATATATCAAGAACTGACCAATCTGGCGTCTCCAGGTCGCTATG 921  
 287 rGValIleValThrGluIleAspProIleCysAlaLeuGlnAlaThrMet 303  
 922 GAAGGCTACAGGTCCGCGCATCGAAGAGAGTCTCAAGATGTGATAT 971  
 304 GluGlyLeuGlnValLeuPheLeuGluAspValValSerGluValAspI 320

972 CTTCGTATACATGACAGAAACTCGATATCATCTCTGTGCATGATGG 1021  
 320 ePheValThrThrThrLysAsnLysAspIleIleMetValAspHisMetAl 337  
 1022 CCCAGATGAAGATMAAGCTATTGTGCGTAACATCGCCACTTCGATTAAC 1071  
 337 rGlyMetLysAsnAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353  
 1072 GAAATTTATACAGATGGCGCTCATGAATCCAGCATCAACACATCC 1121  
 354 GluIleAspMetHisGlyLeuGluThrPheProGlyValLysArgLeuH 370  
 1122 AATCAACCCAGAAATATGACATGTGGAAATTCACATGGCCACGCT..A 1168  
 370 rIleLysProGlnThrAspArgTyrValPheProAspThrAsnSerGlyI 387  
 1169 TCCTCTCTTCTGTGAGGGCGGCTTTTACCTTGGCGGCTACAGGT 1218  
 387 leIleValLeuAlaGluGluValArgLeuMetAsnLeuGlyCysAlaTrpG 403  
 1219 CACCATCTCTGCTTATGTCATGTCAATGCACAAACACAGACTCCCTCA 1268  
 404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaG 420  
 1269 GCTGACCTCTACGAAAGAGA.....GGAATCTCGAGAGAAAGTTT 1312  
 420 nLeuGluLeuThrPasnGluArgSerSerGlyLysGlyLysValT 437  
 1313 ACACACTCCGAGATCTCGATGGAAGTGCCTGCTCCACCTCGGA 1362  
 437 yValLeuProLysHisLeuAspGluLysValAlaIleAlaLeuHisLeuG 453  
 1363 TCTCTGATGTCCACCTTACAAAGCTTACACAGACAGGCTGACTATCAT 1412  
 454 LysPheGlyAlaLysLeuThrLysLeuThrLysAspIleAlaAspTyrI 470  
 1413 CAACTTCCAGTTGAGGGCTTACAAAGTCTGATGCTACCGTTAT 1458  
 470 eTyrValProValGluGlyProTyrLysProAlaHisTyrArgTyr 485  
 seq\_name: sp.plant:Q42939  
 seq\_documentation.block:  
 ID Q42939 PRELIMINARY; PRT: 450 AA.  
 AC Q42939;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE  
 DE HYDROLASE) (ADOCYASE).  
 OS Nicotiana sylvestris (wood tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_Taxid=4096;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RA Mitsui S., Wakasugi T., Sugifura M.;  
 RT "A cDNA encoding the 57 kDa subunit of a cytokinin-binding protein  
 RT complex from tobacco: the subunit has high homology to S-adenosyl-L-  
 RT homocysteine hydrolase.";  
 RL Plant Cell Physiol. 34:1089-1096(1993).  
 CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-  
 CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;  
 CC THEREFORE ADENOSYLHOMOCYSTEINE MAY PLAY A KEY ROLE IN THE  
 CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O - ADENOSINE  
 CC + L-HOMOCYSTEINE.  
 CC -1- COFACTOR: NAD (BY SIMILARITY).  
 CC -1- PATHWAY: ACTIVATED METHYL CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.

DR EMBL: D16139; BAA03710.1; -  
 DR HSP: P10760; 183R.  
 DR InterPro: IPR00043; Ado\_hcyase.  
 DR InterPro: IPR000860; Porphobil\_deam.  
 DR Pfam: PF00670; AdoHcyase; 1.  
 DR PROSITE: PS00738; ADOHCYASE\_1; 1.  
 DR PROSITE: PS00739; ADOHCYASE\_2; 1.  
 DR Hydrolase: NAD; One-carbon metabolism.  
 KW SEQUENCE 450 AA; 49258 MW; B4B010BAF537DADE CRC64;

alignment\_scores:  
 Quality: 1475.00 Length: 459  
 Ratio: 3.965 Gaps: 4  
 Percent Similarity: 81.046 Percent Identity: 63.617

alignment\_block:  
 US-09-759-990-1 x Q42939 ..

Align seg 1/1 to: Q42939 from: 1 to: 450

```

103 ATCCAGATCTTATGTTCTGAGCGTATTCGGCTTACCAT 152
|||||.....|
1 MetProGlyLeuMetAlaCysArgThrGluPheGlyProSerGlnProh 17
153 GAAGGGTTCAGATCTGTTGTTCCCTCCATGACAGTCCAGACGCCG 202
|||||.....|
17 elysGlyAlaLysIleThrGlySerLeuHisMetThrIleGlnThrAlaL 34
203 TCCCTATGAGACATCCAGCTCTGTTGCTGATGTCAGATGGCTTCC 252
|||||.....|
34 AlLeuIleGluThrLeuThrAlaLeuGluAlaGluValArgTrpCysSer 50
253 TGCACATCTCTTACACAGATACAGCCGCTGCTGCTGCTGCTGCTG 302
|||||.....|
51 CysAsnIlePheSerThrGlnAspHisAlaAlaIleAlaIleAlaArg 67
303 CCCAAGGCGCACACAGAGAACGCCGCGTATCCAGTCTTGCGCTGGA 352
|||||.....|
67 PserAlaAla.....ValPheAlaTrpL 75
353 AGGGGAAACATCCAGATCTGAGAGAACATACCGGCTGCTGAC 402
|||||.....|
75 yseGlyIleThrLeuGlnGluTrpTrpCysThrGlnArgAlaLeuAsp 91
403 TGGCCAGATGTCAGAGCCACAGAGTTCGATGATGATGATGATGATG 452
|||||.....|
92 TrpGlyProGlyGlyLeuProAspLeuIleValAspArgGlyLysArg 108
453 TACACTCTCTATCTCCAGAGGCTTC.....GAATTGAAACAG 490
|||||.....|
108 aThrLeuLeuIleHisGluGluValLysAlaGluGluTrpAlaLys 125
491 CCGGTCGCTGTCAGAGCCAGAGTTCAGACATCCAGTACCGGCTGC 540
|||||.....|
125 eglGlyLysLeuProAspProSerSerThrAspAsnAlaGluPheGlnLeu 141
541 GTTCTGTGCTACACTCAAGAGTCTTCAACCAAGTCAAGAACCACTGCA 590
|||||.....|
142 ValLeuThrIleIleArgAspGlyLeuLysThrAspProLeuLysTrh 158
591 CACAGTCTGTCGCCGATGACAGGCTGTTCCGAAGAGACACACAGGTC 640
|||||.....|
158 rLysMetLysGluArgLeuValGlyValSerGluIleThrThrGlyL 175
641 TCCACGGCTTACAGCTGACAGAGGAGGCAAACTCTCTCCAGGC 690
|||||.....|
175 alLysArgLeuLysGlnMetGlnAlaAsnGlyThrLeuLeuPheProAla 191
691 ATCAACGTCAAGCAGCGCTGTACAAAGTCAAGTTCGATACATCTACGG 740
|||||.....|
192 IleAsnValAsnAspSerValThrLysSerLysPheAspAsnLeuTrp 208

```

```

741 CTGCCGCACTCCCTTATCGATGATACACCGTCTCCGATGCATGA 790
|||||.....|
208 YCysArgHisSerLeuProAspGlyLeuMetArgAlaThrAspValMetI 225
791 TCGGGGCAAGACAGCTCGCTCATGGGTTCGCGGATGCGCAAGGCG 840
|||||.....|
225 lAlaGlyLysValAlaLeuValAlaGlyTrpGlyAspAlaLysGly 241
841 TCGCGCAATCCCTCCGTCGCGCAAGGCGCTCGCTTACATCAAGAACT 890
|||||.....|
242 CysAlaAlaAlaLeuLysGlnAlaGlyAlaArgValIleValThrGluL 258
891 CGAACCAATCTCGCTCTCCAGGCTGCCATGGAAGCTACACAGTCCGC 940
|||||.....|
258 eAspProIleCysAlaLeuGlnAlaThrMetGluIleGluGlnValLeu 275
941 GCATGAGGAAGTCGTCAGAGTCTGATCTTCTGATGATGACAGGA 990
|||||.....|
275 hTrLeuGluAspValIleSerAspValAspIlePheValThrThrGly 291
991 AACTGCGATATCATCTCTGTTGATGATGAGCCGCAAGATGAAGG 1040
|||||.....|
292 AsnLysAspIleIleMetValAspHisMetArgLysMetLysAsnAsn 308
1041 TATTCGCGTATCAATCGCCGCTTCGATACGAATGATACAGATGCC 1090
|||||.....|
308 alLeuValCysAsnIleGlyHisPheAspAsnGluLysPheMetLeuGly 325
1091 TCATGAATATCCAGGATCAAGCAGATCCCAATGACGCAATACGAC 1140
|||||.....|
325 euGluThrPheProGlyValLysArgIleThrIleLysProGlnThrAsp 341
1141 ATGTGGAAATTCACAGATGCGCAGCT..ATCCCTCTTCTGCTGAGG 1187
|||||.....|
342 ArgTrpValPheProAspTrpHisAsnSerGlyIleIleValLeuAlaGlu 358
1188 CCGCTTCTTAACTTGGCTGCTGCTACAGGTCACCATCTTCTGATATG 1237
|||||.....|
358 yArgLeuMetAsnLeuGlyCysAlaThrGlyHisProSerPheValMet 375
1238 CAATGTCATCACAAACAGACAGCTGCTGCTGCTGCTGCTGCTGCA 1287
|||||.....|
375 eTrpSerPheThrAsnGlnValIleAlaGlnLeuGluLeuTrpLysGlu 391
1288 AGA.....GGAATCTCGAGAGAGGTTTACACACTCCGAAGCAT 1331
|||||.....|
392 LysSerThrGlyLysTrpGlyLysValGlyValLeuProLysHisLe 408
1332 CGATGAAGAAAGTCTGCTGCTCCACCTGCGATCTCGATGTCACCTTA 1381
|||||.....|
408 uAspGlyLysValAlaAlaLeuHisLeuGlyLysLeuGlyAlaArgLeu 425
1382 CAAGGTTACAGAAAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1431
|||||.....|
425 hTrpLysLeuSerLysAspGlnAlaAspTrpIleSerValProValGlu 441
1432 CTTTACAAAGTCTGATGCTTACCTTAT 1458
|||||.....|
442 ProTyrLysProProHisTyrArgTyr 450

```

seq\_name: sp\_bacteriap:Q9PEJ1

seq\_documentation\_block:

ID Q9PEJ1 PRELIMINARY; PRT; 446 AA.  
 AC Q9PEJ1;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)  
 DE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE  
 DE HYDROLASE) (ADOHCYASE).  
 GN XF1037.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;



323 laLeuAsnAlaLeuAlaGlyValGlnLysIleAsnIleLysProGlnVal 339  
 1138 GACATGTGGGAATTCACGATGGCCACGATCCCTCTCTCTGTCGAGG 1187  
 340 AspLysPheIleLeuProAsnGlnAsnThrLeuPheLeuAlaGluG 356  
 1188 CCGCTCTTACCTTGGCTGCTACAGTTCACCCATCTTTGTTATGT 1237  
 356 YArgLeuValAsnLeuGlyCysAlaThrGlyHisProSerPheValMetS 373  
 1238 CAATGTCTTCACAAACGACACCTGCTCAGTCTCAGCTCTGAGAAAG 1287  
 373 eArAsnSerPheAlaAsnGlnThrLeuAlaGlnIleAspLeuTrpGlnAsn 389  
 1288 AGAGAAATCTCGACAGAGAGTTTACACATCTCCAGACATCTCGATGA 1337  
 390 LysAspValYArgGlnLysAsnValYArgLeuProLysLysLeuAspG 406  
 1338 AGAGTGTGCTGCTCCACCTCGATCTCTGATGTCACCTTCAACAAAG 1387  
 406 uGluValAlaArgLeuHisLeuGlnLysIleGlyValLysLeuThrThrL 423  
 1388 TTACACAGAACGACGAGCTGACTACATCAACGTTCCAGTTGAGGTCCTTAC 1437  
 423 eutHrAlaAsnGlnAlaAlaTyrLeuGlyIleSerValGlnGlyProPhe 439  
 1438 AAGCTGATGCTTACCTTAT 1458  
 440 LysProGlnHisTyrArgTyr 446  
 seq\_name: sp\_bacteriap:Q9ABH0  
 seq\_documentation\_block:  
 ID Q9ABH0 PRELIMINARY; PRT; 463 AA.  
 AC Q9ABH0:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-DEC-2001 (TREMBLrel. 17, Last sequence update)  
 DE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE  
 HYDROLASE) (ADHOCYASE).  
 OS CC0257.  
 NC CAulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 CC CAulobacter.  
 OX NCBI\_taxid=69394;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Uitterback T., Tran K., Wolf A., Vamathavan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of CAulobacter crescentus";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001)  
 CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-  
 ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;  
 CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR  
 CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE  
 CC + L-HOMOCYSTEINE.  
 CC -1- COFACTOR: NAD (BY SIMILARITY).  
 CC -1- PATHWAY: ACTIVATED METHYL CYCLE.  
 CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.  
 DR EMBL: AE005699; AAK22244.1; -.  
 DR HSSP: P10760; 1B3R.  
 DR TIGR: CC0257; -.  
 DR InterPro: IPR000043; Ado\_hcyase.  
 DR Pfam: PF00670; Adhcyase; 1.

DR PROSITE: PS00738; ADHOCYASE\_1; 1.  
 DR PROSITE: PS00739; ADHOCYASE\_2; 1.  
 KW Complete Proteome; Hydrolase; NAD; One-carbon metabolism.  
 SQ SEQUENCE 463 AA; 50760 MW; 0880060406FD9EF9 CRC64;  
 alignment\_scores:  
 Quality: 1374.00 Length: 475  
 Ratio: 3.606 Gaps: 3  
 Percent Similarity: 80.211 Percent Identity: 58.105  
 alignment\_block:  
 US-09-759-990-1 x Q9ABH0 ..  
 Align seg 1/1 to: Q9ABH0 from: 1 to: 463  
 34 GAGTACAGAAATTCGCCGACATCAACCTCCATGTCGCGCGTAAGAACT 83  
 3 AspTyrIleValLysAspIleSerLeuAlaAspPheGlyArgLysGluI 19  
 84 TACCCCTTCGAGAGAGAAATGCCAGGCTTATGCTTCTGTCGAGCGTT 133  
 19 eAlaIleAlaGlnThrGlnMetProGlyLeuMetAlaThrArgAlaGluT 36  
 134 ATTCGCTTCTAAGCCATTGAGAGGCTGTCAGAAATCTCTGTTCCCTCCAC 183  
 36 YGlyProGlnGlnIleLeuLysGlyAlaAlaArgIleAlaGlySerLeuHis 52  
 184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACATCCAGCTCTTGTCG 233  
 53 MetThrIleGlnThrAlaValLeuIleGlnThrLeuThrAlaLeuGlyAl 69  
 234 TGATGTCAGATGGGCTTCTGCAACATCTTCTTACACAAAGATACAGCG 283  
 69 aGluValArgTyrPalaSerCysAsnIlePheSerThrGlnAspHisAla 86  
 284 CTGCTGTATCTGTGTGCGCCCAACAGAGCACACACAGAGAACGCCGCG 333  
 86 IAlaAlaIleAlaIleAla.....AlaGly 93  
 334 ATCCAGTCTTTCGCTGGAAGGCGCAACATCCCAAGATATCGGAGAA 383  
 94 ValProValPheAlaPheLysGlyLysAsnLeuValGluTyrTrpGluTyr 110  
 384 CACATACCGCGCTCTCACATGGCCAGATGTCAGAGGCCACACAGGTTG 433  
 110 rAlaHisLysIlePheGlnThrHisAspGlyTyrProAsnLeuIleL 127  
 434 TCGATGATGGTGGTATGCTACATCTTCATCTCCAAAGGCTTCGAAATTC 483  
 127 euAspAspGlyGlyAspAlaThrLeuLeuCysValLeuGlyProLysAla 143  
 484 GAACAGCGCGGTGCTGTCCAGAGCCCAACAGAAAGCTGACAACTCGAATA 533  
 144 GluLysAspProSerIle.....LeuAsnAsnProGlnAsnLuglu 158  
 534 CCGTGGCTTCTTACATCAACGAGCTTTCACAAAGACAAAGAAC 583  
 158 uGluAlaLeuTyrAlaValMetLysLysTyrLeuAlaGluLysProGlyP 175  
 584 ACTGGCACACAGTGTCTCCGCGCATGAACGCTGTTCCGAGAGACAA 633  
 175 heYrSerAlaIleArgAlaIleGlyGlyValSerGlnGluThrThr 191  
 634 ACAAGTGTCCACCGCTTACAGATGAGAGAGGAGGCAAACTCTCTT 683  
 192 ThrGlyValHisArgLeuTyrGlnMetAlaGlnLysAspGluLeuProH 208  
 684 CCCAGCATCAAGCTCAACGAGCGTGTACAAAGTCAAGTTCGATTAACA 733  
 208 eProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspAsnL 225  
 734 TCTACGCTCGCGCCACTCCCTTATGATGATGATCAACCGTCTCCGAT 783





```

108 uGlnTyrTrpAspTyrThrAspArgIlePheGlnTrpAlaAspGlyLyl 125
419 GCCACAGACAGGTTGTCATGATGGTGGTGGTACACTCTCACTCC 468
125 euserAspMetIleLeuaspAspGlyGlyAspAlaThrMetTyrIleLeu 141
469 AAGGCTTCGAAATTCGAAACACCGCGTGTTCGAGCCACAGAAC 518
142 IleGlyAlaArgAlaGlnIleGlyGlyAspValLeuSeraspProGln 158
519 TGACAACTCGAAATACCGCTGCTTGTACACGACGAGCTTCA 568
158 rGlnIleGlnGlyTyr.....PheTyrAlaGlnValLysLysArgLeu 173
569 ACCAAGACAGAACACACGACACAGATTGCTCCGCGCATGAACGTT 618
173 ysaIleSerProGlyPhePheThrLysGlnLysAlaIleArgGlyVal 189
619 TCCGAGACAGACAAACAGAGTGTCCACCGCTCTACACGCTCGAAGA 668
190 ThrGlnGlnThrThrThrGlyValAsnArgLeuTyrGlnLeuGlnLys 206
669 GGGCAAACTCCCTCTCCACGACATCAAGCTCAAGACGCTGTACAAAG 718
206 sGlyLeuLeuProPheProAlaIleAsnValAsnAspSerValThrLys 223
719 CCAAGTTGATTAATCATTCAGGCTGCCGCACTCCCTTATGATGATC 768
223 eLysPheAspAsnLysTyrGlyLysLysGlnSerLeuValAspGlyLe 239
769 AACCGTCTCCGATGATGATGATGATGATGATGATGATGATGATGAT 818
240 ArgArgGlyLysThrAspThrMetMetAlaGlyLysValAlaValCys 256
819 TTACGCGGATGTCGCGAAGGCTGCTCAATCCCTCGTGCGCAAGCG 868
256 yTyrGlyAspValGlyLysSerSerAlaSerLeuLysGlyAlaGly 273
869 CTTCGCTTATCATCAGAACTCGAACCAATCTGGCTCTCCAGGCTG 918
273 IaArgValLysValThrGlnValAspProIleCysAlaLeuGlnAla 289
919 ATGAAGGCTACAGAGTCCCGCGCATCAGAGAGTGTGATGATGATG 968
290 MetAspLysPheGlnValValThrLeuGlnAspAlaAlaProThrAla 306
969 TATCTTCTTACATGACAGAGAACTGGATATCATCTCTGTGACATG 1018
306 pIleValIleThrThrThrGlyAsnLysAspValValThrLeuAspHis 323
1019 TGGCGCCGATGAGATGAGGCTATTGTGCTGATACGCGCACTTCAT 1068
323 eLysSerMetLysAspMetValIleValGlyAsnIleGlyHisPheAsp 339
1069 AACGAAATGATACAGATGCGCTCATGAATATCCAGGCAATCAGCA 1118
340 AsnGlnIleGlnValAlaSerLeuArg.....AsnLeuLysTrp 353
1119 CCCAATCAAGCGCAATACAGATGAGGATTCACAGATGCGCAGCT 1168
353 rAsnValLysProGlnValAspMetIleThrPheProAspGlyLysArg 370
1169 TCTCTCTTCTTGTGAGGCGCGCTTCTTAACCTGGCTGCGCTACAG 1218
370 eTlleLeuLeuSerGlnGlyArgLeuLeuAsnLeuGlyAsnAlaThrGly 386
1219 CACCATCTTCTTGTATGATGATGATGATGATGATGATGATGATGAT 1268
387 HisProSerPheValMetSerAlaSerPheThrAsnGlnValLeuAla 403
1269 GCTGACCTCTACGAAAGAGAGAAATCTCGAAGAGAGAGAGAGAG 1318
403 nIleGlnLeuPheThrLysGlyGlnGlnTyrGlnAsnGlnValTyrVal 420

```

```

1319 TTCGAGACATCTCATGAGAGAGTGGCTGCCCTCCACCTGCAATCTC 1368
420 euProLysHisLeuaspGlnLysValAlaArgLeuHisLeuaspLysLeu 436
1369 GATGTCACCTTACAAAGCTTACACAGAGAGGCTGACTATCACTAC 1418
437 GlyAlaArgLeuThrGlnLeuSerGlyGlnIleAlaIleTyrIleGly 453
1419 TCCAGTTGAGGCTCTTACAAAGTGTGATGCTTACCGTTAT 1458
453 LThrProGlnGlyProPheLysProGlnHisTyrArgTyr 466

```

seq\_name: sp\_bacteria:092NMS

```

seq_documentation_block:
ID 092NMS PRELIMINARY; PRT; 462 AA.
AC 092NMS;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE
DE HYDROLASE) (ADOHCYASE).
GN AHCY.
OS Erythrobacter sp. (strain OCH 114).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Roseobacter.
OX NCBI_TaxID=2434;
RN [1]
RP SEQUENCE FROM N.A.
RA Nishimura K., Shimada H., Shimmen T., Obayashi T., Masuda T., Ohta H.,
RA Takamiya K.;
RT "Photosynthetic regulatory gene cluster in an aerobic photosynthetic
RT bacterium, Roseobacter denitrificans."
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS.
CC THEREFORE ADENOSYLHOMOCYSTEINE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE
CC + L-HOMOCYSTEINE.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
CC EMBL: AB020211; BA34645.1; -.
DR HSSP: P10760; IBSR.
DR InterPro: IPR000043; ADO_hcyase.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF00670; ADOHCYase; 1.
DR PROSITE: PS00738; ADOHCYASE_1; 1.
DR PROSITE: PS00739; ADOHCYASE_2; 1.
DR Hydrolase: NAD; One-carbon metabolism.
KW SEQUENCE 462 AA; 50484 MW; 8DBD3A6F3F896FBE CRC64;

```

alignment\_scores: Quality: 1337.00 Length: 477

Percent Similarity: 77.358 Percent Identity: 58.281

alignment\_block: US-09-759-990-1 x 092NMS

Align seg 1/1 to: 092NMS from: 1 to: 462

```

34 GAGTACAGAAATTCGCGACATCAACCTCTGCGCGGTAAGAACT 83
4 AspPheIleValLysAspIleAlaLeuAlaGlnPheGlyArgLysGln 20
84 TACCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 133
20 unspleAlaGlnThrGlnMetProGlyLeuMetAlaLeuArgAlaGln 37

```



```

134 ATTCGGCTTCTAAGCCATTGAAGGCTGTCAGAAATCTGCTGCTCCAC 183
137 YIYLASPSerLysProLeuAlaArgIleValGlySerLeuHis 53
184 ATGACAGTTCAGACAGCGCTCCTCATTCGACAGACATCAACCTTTGTGC 233
54 MethIleIleIntrAlaValLeuIleGluThrLeuValAlaLeuGlyAl 70
234 TGATGTAGATGGGCTTCTGCAACATCTTCTACACAAATACAGCCG 283
70 aasPValArgTTrpAlaSerGlyAsnIlePheSerThrGlnAspHisAla 87
284 CTCTGCTATTCGTTGTGGCCCAACAGCACACAGAGAACGCCCGGT 333
87 laIaIaIaIeIaIaIaGly.....Gly 94
334 ATCCAGTCTTCCGCTGGAAGGGGAAACATCCAGATACAGGAGAA 383
95 ThrProValPheAlaIleLysGlyGlnSerLeuGluGluHisTrpAsp 111
384 CACATACCGGCTCTCATATGGCAGATGTCAGGCCCAACAGAGTTG 433
111 rLeuAspArgSerPheMetPheGluAsp.....GlyProAsnLeuIle 126
434 TCGATGATGGTGTGATGCTACATCTCCATCTCCAAGGCTTCAATTC 483
126 euAspAspGlyGlyAspAlaThrLeuTrpValLeuLeuGlyAlaArgAla 142
484 GAACACCGCGT.....GCTGTCCAGAGCCACAGACGTCGACAAAC 527
143 Glu...AlaGlyGluGluIleIleProValProThrSerGlu..... 155
528 CGAATACCGCTGCTTCTGCTACATCTCAACAGGCTTCAACAGACA 577
156 GluGluGluAlaIleLysAlaGlnIleLysLysArgMetAlaAlaSer 172
578 AGAACACGCGGACACAGTTCCTCCGACATGAGACGCTTCCGAGAG 627
172 roGlyTrpPheThrLysValArgAspGlnIleLysGlyValSerGluGlu 188
628 ACAACAACAGGTCGACCGCTTACACGCTCGAAGAGGAGGCAAACT 677
189 ThrThrThrGlyAlaHisArgLeuTrpAspLeuValLysGlnGlyGlu 205
678 CCTTCTCCACGCTCAAGCTCAACGAGCGCTTCAAAAGTCCAAAGT 727
205 uProPheProAlaIleAsnValAsnAspSerValThrLysSerLysPhe 222
728 ATAACATCTACGCTGCGGCGACCTTCATCGANTGATCAACCGTGT 777
222 spAsnLysTrpGlyCysLysGlnSerLeuValAspGlyIleArgArgAla 238
778 TCCGATGTCATGATCGCGGCAAGACGCTCTCATGGGTTACGGGCA 827
239 ThrAspThrMetAlaIleLysValAlaValAlaMetGlyTrpGlyAs 255
828 TGTGCGCAAGGCTGCGCTCAATCCCTCCGTGGCCAGCGCTCGCTTA 877
255 PValGlyLysGlySerAlaLeuAlaSerLeuArgGlyAlaGlyAlaArgVal 272
878 TCATCAGAACTCGACCAATCTGCGCTCCAGGCTCGCATGGAAGGC 927
272 ysValThrGluValAspProIleCysAlaLeuGlnAlaAlaMetAspGly 288
928 TACCAAGTCCGCGCGATCGAAGATGTCAGATGTCATCTTCTGCT 977
289 PheGluValAlaLeuLeuGlnAspValAlaGlySerAlaAspIlePhe 305
978 TACATGCACAGAACTGGATATCATCTGTTGACATGATGGCCGCA 1027
305 eThrThrThrGlyAsnLysAspValIleArgIleGluHisMetArgAla 322

```

```

1028 TGAAGATTAAGCTATTTGCTGCTAATCGGCGCACTTGATTAAGAAAT 1077
322 eLysAspMetAlaIleValGlyAsnIleGlyHisPheAspAsnGluIle 338
1078 GATACAGATGGCTTCATTAATATACCGAGCATCAACATCCCAATCAA 1127
339 GlnValAlaAlaLeuLysAsnHis.....LysTrpThrAsnIleLys 352
1128 GCCGAATTAAGCAATGTGGAAATCCCAATGCGCACGCTATCTCTTC 1177
352 sGluGlnValAspMetIleGluMetProAsnGlnValAlaArgLeuIleLeu 369
1178 TTGCTGAGGGCGGCTTCTTAACCTTGGCGCTGACGCTCAAGCT 1227
369 euSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 385
1228 TTGCTTATGTCATATGTCATTCACAAACAGACACTGCTCAGCT 1277
386 PheValMetSerAlaSerPheThrAsnGlnValAlaGlnIleGluLe 402
1278 CTACGAAAGAGAGAAATCTCGAGAGAAGGTTTACACTCCAGAC 1327
402 uTrpThrArgAlaAspAlaTrpAspAsnGluValTrpIleLeuProLysH 419
1328 ATCTCGATGAAGAATCGCTCGCTCCACCTCGATCTCGATGCCAC 1377
419 lIleuAspGluLysValAlaAlaArgLeuHisLeuAspArgIleGlyValLys 435
1378 CTACAAAGCTTACACAGAACAGGCTGACTCATCAACAGCTTCCAGTTGA 1427
436 LeuThrProLeuAspProGluGlnAlaAlaTrpIleGlyValLysProG 452
1428 GGGTCTTACAGTCTGATGCTTACCGTTAT 1458
452 uGlyProPheLysProGluHisTrpArgTrp 462

```

```

seq.name: SP_rodent:Q91WF1
seq_documentation_block:
ID Q91WF1 PRELIMINARY; PRT; 432 AA.
AC Q91WF1:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE S-ADENOSYLHOMOCYSTEINE HYDROLASE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (CCN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015304; AAH15304.1; -.
KW Hydrolase.
SQ SEQUENCE 432 AA; 47688 MW; 15CCDD20B7088D5CA CRC64;

```

```

alignment_scores:
Quality: 1279.00 Length: 478
Ratio: 3.665 Gaps: 5
Percent Similarity: 73.013 Percent Identity: 56.485

```

alignment\_block:

US-09-759-990-1 x Q91WF1

Align seg 1/1 to: Q91WF1 from: 1 to: 432

```

37 TACAGAAATGCGACATCAACTTCATGTTCTCGCGCGTAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 TyrLysValAlaAspIleGlyLeuAlaIleArgPheGlyArgLysAlaLeuAs 23

```

```

87 CCTTGTGAGAGAAATGCCAGGTCTATGTTCTTCGTAGCGTTAT 136
   ::::::::::::::::::::::::::::::::::::::::::::::
23 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgGluMetTyr 40
   ::::::::::::::::::::::::::::::::::::::::::::::
137 CGGCTTTAGAGCATGGAAGGGTGTGAGAAATCTGTGCTCCCTCCACATG 186
   ::::::::::::::::::::::::::::::::::::::::::::::
40 eAlaSerLysProLeuLysGlyAlaArgGluLeuGlyCysLeuHisMet 56
   ::::::::::::::::::::::::::::::::::::::::::::::
187 ACAGTCCAGACAGCGCTCTCATCGAGACATCCACAGCTTGTGTGTGA 236
   ::::::::::::::::::::::::::::::::::::::::::::::
57 ThrValGluThrAlaValLeuIleGluThrLeuValAlaLeuGluAlaG 73
   ::::::::::::::::::::::::::::::::::::::::::::::
237 TGTCAAGATGGGCTTCTGCAACATCTTCTACACAGATATACAGCGGTG 286
   ::::::::::::::::::::::::::::::::::::::::::::::
73 uValArgPheSerSerCysAsnIlePheSerThrGlnAspHisAlaAla 90
   ::::::::::::::::::::::::::::::::::::::::::::::
287 CTGTATGTTGTTCGGCCCAACAGGCAACACGACGAAAGCCAGCGGTATC 336
   ::::::::::::::::::::::::::::::::::::::::::::::
90 lAlaIleAla.....LysAlaGlyLe 97
   ::::::::::::::::::::::::::::::::::::::::::::::
337 CCAGTCTTGCCCTGGAAAGGCGAAACATCCCAAGATCTGGGAGAACAC 386
   ::::::::::::::::::::::::::::::::::::::::::::::
98 ProValPheAlaTrpLysGlyGluThrAspGluGluTyrLeuTrpCysI 114
   ::::::::::::::::::::::::::::::::::::::::::::::
387 ATACCGCGCTCACATGGCCACATGGTTCAGGCCCAAGCAGCAGGTGTGC 436
   ::::::::::::::::::::::::::::::::::::::::::::::
114 eGluGlnThrLeuHisPheLysAspGlyPro..LeuAsnMetIleLeuA 130
   ::::::::::::::::::::::::::::::::::::::::::::::
437 ATGATGGTGTGATGTACATCTCTCATCTCCAAAGGCTTCGAATTCGA 486
   ::::::::::::::::::::::::::::::::::::::::::::::
130 sPaspGlyGlyAspLeuThrAsnLeuIle..... 139
   ::::::::::::::::::::::::::::::::::::::::::::::
487 ACAGCGGCTCTGTCCAGAGCCAAACAGAACTGACAACTCGAATACCG 536
   ::::::::::::::::::::::::::::::::::::::::::::::
139 ..... 139
   ::::::::::::::::::::::::::::::::::::::::::::::
537 CTGCGTCTTGTCTACACTCAAGCAGGTCTTCAACCAAGACAAACCACT 586
   ::::::::::::::::::::::::::::::::::::::::::::::
139 ..... 139
   ::::::::::::::::::::::::::::::::::::::::::::::
587 GGCACACA.....GTGTGCGCGGCATGACGGTGTTCGGA 624
   ::::::::::::::::::::::::::::::::::::::::::::::
140 ..HisThrLysTyrProGlnLeuLeuSerGlyIleArgGlyIleSerGlu 155
   ::::::::::::::::::::::::::::::::::::::::::::::
625 GAGCAACAACAGGTGTCCACCGGCTTACACAGCTCGAAGAGAGGCA 674
   ::::::::::::::::::::::::::::::::::::::::::::::
156 GluThrThrThrGlyValHisAsnLeuTyrLysMetMetSerAsnGlyT 172
   ::::::::::::::::::::::::::::::::::::::::::::::
675 ACTTCTCTTCCAGCCATCAAGTTCACAGCAGCTGTTCACAAAGTCCAAGT 724
   ::::::::::::::::::::::::::::::::::::::::::::::
172 eLeuLysValProAlaIleAsnValAsnAspSerValThrLysSerLysP 189
   ::::::::::::::::::::::::::::::::::::::::::::::
725 TCGATACATCTAGGCTGCCGCCATCCCTTATGATGATGATCAACCGT 774
   ::::::::::::::::::::::::::::::::::::::::::::::
189 heAspAsnLeuTyrGlyCysArgGluSerLeuIleAspGlyIleLysArg 205
   ::::::::::::::::::::::::::::::::::::::::::::::
775 GCTTCGATCATGATCGCGGCAAGACAGCTCTGTCATGGGTTACG 824
   ::::::::::::::::::::::::::::::::::::::::::::::
206 AlaThrAspValMetIleAlaGlyLysValAlaValAlaIleAlaGlyT 222
   ::::::::::::::::::::::::::::::::::::::::::::::
825 CGATGTGGCAAGGCTGCGCTCAATCCCTCGTGGCCCAAGGCGCTCGCG 874
   ::::::::::::::::::::::::::::::::::::::::::::::
222 yAspValGlyLysGlyCysAlaGlnAlaLeuArgGlyPheGlyAlaArgV 239
   ::::::::::::::::::::::::::::::::::::::::::::::
875 TTATCATCACAGAACTGCAACCAATCTGCGTCTCAGGCTCCATGAA 924
   ::::::::::::::::::::::::::::::::::::::::::::::
239 aIleIleThrGluIleAspProIleAsnAlaLeuGlnAlaIleMetGlu 255
   ::::::::::::::::::::::::::::::::::::::::::::::
925 GGCATCAGAGTCCGCCGCATCGAGGAAGTCTCAAGGATGTCGATATCTT 974
   ::::::::::::::::::::::::::::::::::::::::::::::
256 GluTyrGluValThrThrMetAspGluAlaCysLysGluGlyAsnIlePhe 272
   ::::::::::::::::::::::::::::::::::::::::::::::
975 CGTTACATGCACAGAACTCGATATCATCTCTGTGACATGATGGCC 1024

```

```

   ::::::::::::::::::::::::::::::::::::::::::::::
272 eValThrThrThrGlyCysValAspIleIleLeuGlyArgHisPheGluG 289
   ::::::::::::::::::::::::::::::::::::::::::::::
1025 AGATGAAGATPAGGCTATTGTCCGTAACATCGCCACATTCGATACGA 1074
   ::::::::::::::::::::::::::::::::::::::::::::::
289 lMetLysAspAspAlaIleValCysAsnIleGlyHisPheAspValAsn 305
   ::::::::::::::::::::::::::::::::::::::::::::::
1075 ATTGATACAGATGGCTCATGAATATCCAGGATCAACAGCATATCCCAAT 1124
   ::::::::::::::::::::::::::::::::::::::::::::::
306 lIleAspValLysTrpPheAsnGlu...AsnAlaValGluLysValAsn 321
   ::::::::::::::::::::::::::::::::::::::::::::::
321 eLysProGluValAspArgTyrTrpLeuLysAsnGlyArgArgIleIle 338
   ::::::::::::::::::::::::::::::::::::::::::::::
1175 TTCTGTGAGGGCGCCCTTCTTAACTTGGTGGGTACAGTCAACCA 1224
   ::::::::::::::::::::::::::::::::::::::::::::::
338 eLeuAlaGluGlyArgLeuValAsnLeuGlyCysAlaMetGlyHisPro 354
   ::::::::::::::::::::::::::::::::::::::::::::::
1225 TCTTTCGTTATGTCAATGTCAATTCACAAACACAGACTGCTACGTCGA 1274
   ::::::::::::::::::::::::::::::::::::::::::::::
355 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIleG 371
   ::::::::::::::::::::::::::::::::::::::::::::::
1275 CCTTACGAAAGAGAGGAATCTGACAGAAAGTTTACACTTCCGA 1324
   ::::::::::::::::::::::::::::::::::::::::::::::
371 uLeuTrpThrHisProAspLysTyrProValGluValHisPheLeuPro 388
   ::::::::::::::::::::::::::::::::::::::::::::::
1325 AGCATCTCGATGAGAAAGTGGCTGCTCCACTGCTCGATCTCGATGTC 1374
   ::::::::::::::::::::::::::::::::::::::::::::::
388 yLysLysLeuAspGluAlaValAlaGluAlaHisLeuGlyLysLeuAsnVal 404
   ::::::::::::::::::::::::::::::::::::::::::::::
1375 CACCTTACAAAGCTTACACAGAGCAGGCTGACTATCAAGCTTCACAGT 1424
   ::::::::::::::::::::::::::::::::::::::::::::::
405 LysLeuThrLysLeuThrGluLysGlnAlaGlnTyrLeuGlyMetProI 421
   ::::::::::::::::::::::::::::::::::::::::::::::
1425 TGAGGCTCTTACAACTGATGCTTACCGTAT 1458
   ::::::::::::::::::::::::::::::::::::::::::::::
421 eAsnGlyProPheLysProAspHisTyrArgTyr 432

```







```

63 MetThrIleGlnThrAlaValLeuIleGlnThrLeuThrAlaLeuGlyPr 79
224 TGATTCAGATGGGCTTCTGCAACATCTCTACACAAAGATACACCG 283
79 GGIValAlaArgTrpCysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATCGTGTGCGCCCAACAGACGACACAGAGAAAGCCGCGT 333
96 IAlaIAlaIleAlaArgAspSerAlaSer..... 105
334 ATCCAGTCTTGCTGCGAAGGCGAAACACTCCCAAAATACTGGAGAA 383
106 .....ValPheAlaIleTrpLysGlyGlnThrLeuGlnGlnIleTrpTrpCys 120
384 CACATACCGCGCTTCACATGGCCAGATGGTCAAGCGCCACAGACAGGTG 433
120 sThrGlnAlaIleAlaLeuAspTrpGlyGlyGlyProAspLeuIleV 137
434 TCGATGATGGTGGTGTGATGCTACACTCTCATCTCCAAAGGCTTC..... 477
137 AlAspAspGlyGlyAspThrThrLeuLeuIleHisGlnGlyValIysAla 153
478 .....GAATTGAAACAGCGCGTGTCCAGAGCCAAACAGAGCTGA 521
154 GlnGlnGlnIleTrpGlnIleLysMetProAspProAlaSerThrAs 170
522 CAACCTCGAATACCGCTGCTTGTGTACATCAACAGAGCTTCAAC 571
170 pAsnAlaGlnPheGlnIleValLeuThrIleIleArgAspGlyLeuIysV 187
572 MAGACAAGACCACTGGCACACAGTTGCTCGCGCATGACGAGTTC 621
187 AlAspProThrLysTrpArgLysMetLysAspArgIleValGlyValSer 203
622 GAAGAGACAAACAGAGTGTCCACCGCTTACACGCTCGAAGAGGAGG 671
204 GlnGlnIleThrThrThrGlyValLysArgLeuIleGlnMetGlnAlaAspAs 220
672 CAACCTCTCTTCCAGCCATCAACGTCACAGAGCTGTTCACAAAGTCA 721
220 nSerLeuLeuPheProAlaIleAsnValaAsnAspSerValThrLysSerL 237
722 AGTTCGATAACATCTACGCGTCCCGCCACTCCCTTATCGATGATCAAC 771
237 yAspPheAspAsnLeuTrpGlyCysArgHisSerLeuProAspGlyLeuMet 253
772 CGGCTCGATGATCATGATCGCGGCAAGACGCTCGCATGGGTTA 821
254 ArgAlaIleThrAspValMetIleAlaGlyLysValAlaValaValCysGlyTrp 270
822 CGGCGATGTCGGCAAGGCTGCGCTCAATCCCTCGTGGCCAAAGGCGCTC 871
270 rGlnAspValGlyGlnGlyCysAlaIleAlaLeuLysGlnAlaGlyAla 287
872 GCGTTATCATCAGAACTGACACCATTCGCGCTCCAGGCTCCCATG 921
287 rGValIleValIleThrGlnIleAspProIleCysAlaLeuGlnAlaLeuMet 303
922 GAAGGCTACAGAGTCCCGCATCGAGAAAGTCTCAAGATGTCGATAT 971
304 GlnGlyLeuGlnIleValLeuThrLeuGlnAspValaValSerGlnAlaAspI 320
972 CTTCGTTACATGACAGAAACTCGCATATCATCTGTGTCAGATGATGG 1021
320 ePheValIleThrThrThrGlnLysAsnLysAspIleIleMetLeuAspHisMet 337
1022 CCCAGATGAAGGATAGGCTATTTGCTGATACATCGGCGCATCATAC 1071
337 rGlySerMetLysAsnAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353
1072 GAATATGATACAGATGGCTCATGAATATCCAGGACATCAAGACATCCC 1121
354 GlnIleAspMetLeuGlyLeuGlnIleThrTrpProGlyIleLysArgIleHis 370

```

```

1122 AATCAAGCCAGAAATACGACATGTGGAAATTCACAGAGGCCAC...GCTA 1168
370 rIleLysProGlnIleThrAspArgTrpValPheProGlnIleThrAsnThrGlyI 387
1169 TCCTCTCTCTTGTGCTGAGGCGCGCTTCTTAACCTTGGCTGGCTACAGT 1218
387 lIleValIleLeuAlaGlnGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
1219 CACCATCTTCTGTTATGTCAATGTCAATTCACAAACAGACACTGCTCA 1268
404 HisProSerPheValMetSerCysSerPheThrAsnGlnIleAlaGln 420
1269 GCTGACCTTACGAAAAAGAA.....GGAATCTCGAAGAAAGGTTT 1312
420 nLeuGlnLeuTrpAsnGlnLysAlaSerGlyLysTrpGlyLysValT 437
1313 ACACACTTCGGAAGCATCTGATGAGAAAGCGCTCGGCTCCACTCGGA 1362
437 yValLeuProLysHisLeuAspGlnLysValaIleAlaLeuHisLeuGly 453
1363 TCTCTGATGTCCACTTACAAAGCTTACACAGAGAGGCTGACTACAT 1412
454 LysLeuGlyAlaLysLeuThrLysLeuSerProSerGlnAlaAspTrpI 470
1413 CAACGTTCCAGTTGAGGCTCTTACAAAGTCTGATCTTACGCTTAT 1458
470 eSerValProIleGlnGlyProTrpLysProProHisTrpArgTrp 485
seq_name: /cgn2_6/protodata/2/1aa/5b_comb.pep:US-08-896-005-5
seq_documentation_block:
; Sequence 5, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0337 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:

```

LIBRARY: GenBank  
CLONE: 904132  
US-08-896-005-5

alignment\_scores:  
Quality: 1278.00 Length: 478  
Ratio: 3.662 Gaps: 5  
Percent Similarity: 73.013 Percent Identity: 56.485

alignment\_block:  
US-09-759-990-1 x US-08-896-005-5

Align seg 1/1 to: US-08-896-005-5 from: 1 to: 432

```

37  TACAGATTCGCGACATCAACCTCCAGTTCGCGCGGTAAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7  TATLVAALAAAPRIAGLYLEUALAALATPGLATGGLYSALALEUAS 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87  CCTTGTGAGAGAAATGCGAGCTTATGCTCTTGTGAGCGTTAT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23  PILEAALAGLUSAGLUMETPROGLYLEUMETARGMETARGLUMETYS 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137  CGGCTTCTAAGCCATGAAAGGAGTGTCAAGAACTCTGTTCCCTCAATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40  ERLASERLYSPROLEULYSGLYALATPGLIAGLYCYSEUHIHMET 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187  AAGATTCACACAGCCGCTTCATGAGACACATCAAGCTTGTGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57  TPLVALGLTHTHALVALLEULIAGLUTHTLEUVALALEUGLYALAGL 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237  TGTCAAGATGGGCTTCTGCAACATCTTCTACACAAAGATACAGCGGTG 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73  VALATPGLTPTSERSECYASNLIEPESERTHTGLASPHISALALAA 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287  CTGTATCTGTGTGGCCCAACAGGCAACACAGAAAGCCAGCCGGTATC 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90  IAAALALEUA.....LysALGLYILE 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337  CCAGTCTTGGCGTGAAGGGCGAAACATCTCCGAAATACCTGGAGAACAC 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
98  PROVALPHEALATPGLYSGLYUTHTASPGILUTHTLEUTRPSYI 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387  ATACGCGCTTCACATGGCCAGAGGTGTCAAGGCCACAGACAGGTTCG 436
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
114  EGLUTHTHTLEUHIHSPHELYASPRILEU...LEUASMETILEUA 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
437  ATGATGGTGTGATGCTACACTCTCATCTTCAAGGGCTTCAATTGGA 486
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130  SPASPGLYGLYASPLEUTHTASPLEU...LysALGLYILE 139
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
487  ACAGCGGTGTGTTCAGAGGCAACAGAACTGACAACTCGAATACCG 536
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139  ..... 139
537  CTGCGTCTTGTCTACATCAACAGGTCTTCAACAAAGAACACCACT 586
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139  ..... 139
587  GGCACACA.....GTTCGTCGCGCATGAACGGTGTTCGGA 624
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
140  ..HISTHTLYTPROGLINLEUASERGLYLEARGGLYILESERGL 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
625  GAGACAAACAGAGTGTCCACCGCTTACAGCTGAGAGAGAGGCA 674
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156  GLUTHTHTHTHTGLYALHISASNLLEUTRYLSKEMETSERASNGLY 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
675  ACTCTCTTCCAGCATCAACGTCAACGAGCCTGTACAAATCCAAAT 724
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
172  ELEASVALATPGLIAGLYLEUASVALASPSERVALTHTLYSERLYSP 189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
725  TCGATACATCTAGCGTGGCGCACTCCCTATGATGATCAACCGT 774
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

189  heASPASNLLEUTRYGLYCYASARGGLUSERILEASPRGLYILEYSLARG 205
775  GCTTCCGATGTCATGATCGCGCAAGACAGCTCTGTCATGGGTACGG 824
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206  ALATHTASPRVALMETILEALAGLYLSVALALVALVALALAGLYTRYGL 222
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
825  CGATGTCGCAAGGGCTGCGCTCAATCCCTCGCTGGCAAGGCGCTGCG 874
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222  YASPRALGLYLSGLYCYASALAGLIALLEUATARGLYALARGV 239
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
875  TTATCATCACAGAACTGACCAATCTGCGCTTCCAGCTGCATGGA 924
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239  ALILEITHTHTHTHTLEASPRILEASNALALEUGLIALALEMETGL 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
925  GGTATCCAGTCCCGCGCATCGAGAGAGTCTGATAGAGTGCATATCTT 974
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256  GLYTRYGLVALTHTHTMETASPRGLIALCYASGLYGLULYSNLIEPH 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
975  CGTTACATGACAGGAAACTGCGATTCATCTCTGTGACATGATGCGCC 1024
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272  EVALTHTHTHTHTHTGLYCYVALASPRILEILEUGLYATHTHTHTHT 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1025  AGATGAGGATPAGGCTATGTCGTTACATGCGCCCATTCGATTAAGAA 1074
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289  IMETLYASPRALALEVALCYASNLIEGLYHISPHASPRVALGL 305
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1075  ATTGATACAGATGGCTCATGAATACCAAGCATCAACACATCCCAAT 1124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306  ILEASPRALYSTRPLEASNL...ASNALVALGLULYSVALSNIL 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1125  CAAGCCAGAAATACGATGTGGAAATCCCAATGAGCCAGCATCTCC 1174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
321  ELYSPROGLINVALASPARGYTRYTPLEULYSANGLYALGATGILIEL 338
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1175  TTTCTGTGAGGCGCGCTTCTTAACCTTGGCTGCGCTACAGTCAACCA 1224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
338  EULEUALAGLGLYATGLENVALSNLEUGLYCYASALAMETGLYHISPRO 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1225  TCTTTCGTTATGTCATGTCACAAACGACAGACAGCTGCGTCAAGTCA 1274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355  SERPHEVALMETSERASNSERTHTASNGLVALMETALAGLIEGL 371
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1275  CCTTACGAAAGAGAGAAATCTGAGAGAGGTTTACACACTTCCGA 1324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
371  ULUTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHT 388
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1325  AGCATTCGATGAGAGAGTCCGCTGCGCTCAACCTGATCTGATGTC 1374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388  YSLYSLASPRGLIALVALALAGLIALHISLEUGLYLYSLASVAL 404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1375  CAACCTTCAAGAGTTTACAGAGAGGCTGACTATCAACAGCTTCCAGT 1424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
405  LYSLEUTHTRYLSLEUTHTRYGLINLALAGLIALINTRYLEUGLYMETPRO 421
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1425  TGAAGGCTTACAGTCTGATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421  EASNGLYPROPHELYSPASPRHISTYATGYT 432
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq\_name: /cgn2\_6/plodata/2/1aa/5b.COMB.pep:US-08-896-005-4

seq\_documentation\_block:

Sequence 4, Application US/08896005

Patent No. 5854023

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Lal, Preeti

APPLICANT: Shah, Purnvi

TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/896,005  
FILING DATE: Filed Herewith  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 178277  
US-08-896-005-4

alignment\_scores:  
Quality: 1255.00 Length: 478  
Ratio: 3.617 Gaps: 5  
Percent Similarity: 72.594 Percent Identity: 55.439

alignment\_block:  
US-09-759-990-1 x US-08-896-005-4 ..

Align seg 1/1 to: US-08-896-005-4 from: 1 to: 432

```

37 TACAGAAATGGCGACATCCATCCATGCTCGCGCGTAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7 TTYLysValAlaAspIleGlyLeuAlaIleTyrPheArgLysAlaLeuAs 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
87 CCTTGCTGAGAAAGAAATGCCAGTCTTATGGTCTTCTGAGCGTTATT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
23 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgLysIleTyr 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
137 CGCTTCTTAAGCATTAAGGCTTCAGAAATCTGTTCCGTCCTCCACATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
40 eValAserLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
187 ACAGTCACAGACGCGCTCATGAGACATCACTCAGCTCTTGGTGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
57 ThrValGluThrAlaValIleuIleGluThrLeuValThrLeuGlyAla 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 TGTCAAGTGGGCTTCTGCAACATCTTCTACACAAATACAGCGCGTG 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
73 uValGlnTyrSerSerCysAsnIlePheSerThrGlnAsnHisAlaAla 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
287 CTGCTATCGTGTGTGGGCCAACAGGACACACAGAGAACAGCGCGTATC 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
90 IAlaIleLea.....LysAlaGlyIle 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
337 CCAAGTCTCCGCTGGAAGGCGCAACACTCCAGAAATACCTGGAGAACAC 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
98 ProValIleTyrIleTyrLysGlyIleThrAspGluIleTyrLeuTyrP 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
387 ATACCGCGCTCTCAATGGCCAGATGATGTCAGAGCCACAGCAGGTGTG 436

```

```

114 egluGlnThrLeuTyrPheLysAspGlyPro...LeuAsnMetIleLeuA 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
437 ATGATGGTGGTATGCTACACATCCATCCATCCAGAGCGCTCGAATTGCA 486
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
130 spAspGlyGlyAspLeuThrAsnLeuIle..... 139
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
487 ACAGCGGCTGTCTCCAGAGCCACAGAGAAAGTGCACAACTCGAATAACG 536
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
139 ..... 139
537 CTGCGTCTTGTCTACACTCAAGCAGGCTTTCACCAAGACAAAGAACACT 586
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
139 ..... 139
587 GGCACACA.....GTTCGTCGCGCATGAACGGTGTTCGGA 624
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
140 ...HisThrLysTyrProGlnLeuLeuProGlyIleAlaGlyIleSerGlu 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
625 GAGACAAACAAGAGTGTCCACCGCTCTACACGCTCGAAGAGAGGCA 674
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
156 GluThrThrThrGlyValHisAsnLeuTyrLysMetAlaAsnGlyIle 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
675 ACTCTCTTCCCGCATCAACGTCACAGAGCGCTTACAAAGTCCAGT 724
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
172 eLeuLysValProAlaIleAsnValAsnAspSerValThrLysSerLysP 189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
725 TCGATACATCTACGCGTCCGCGCATCTCTTATCGATGATCAACGCT 774
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
189 heAspAsnLeuTyrGlyCysArgLysSerLeuIleAspGlyIleLysArg 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
775 GCTTCGATGTCATGATCGCGCGCAAGACACTCTCGTCATGGTTACG 824
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
206 AlaThrAspValMetIleAlaGlyLysValAlaValAlaIleGlyTyl 222
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
825 CGATGTCGCAAGGCGTCCGCTCAATCCCTCGTGGCCAGGCGCTCGG 874
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
222 yAspValGlyLysGlyCysAlaGlnAlaLeuArgGlyHeGlyAlaArgV 239
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
875 TTTATCATCAGACAGTCCAGCCCAATCTGGCGTCTCCAGCGTCCATGAA 924
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
239 alIleIleThrGlnIleAspProIleAsnAlaLeuGlnAlaIleMetGlu 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
925 GCGTACAGAGTCCCGCATGAGAGAACTGTCAGAGATGATGATATTCT 974
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
256 GlyTyrGluValThrThrMetAspGluAlaCysGlnGluGlyAsnIlePh 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
975 CGTACATGTCACAGAAACTGCAATCATCTCTGTTGACATGATGGCC 1024
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
272 eValThrThrThrGlyCysIleAspIleIleLeuGlyArgHisPheGlu 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1025 AGATGAAGATTAAGCTATTGTCGTTAACATGGCCACTTCGATACGAA 1074
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
289 IMetLysAspAspAlaIleValCysAsnIleGlyHisPheAspValGlu 305
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1075 ATTGATACAGATGGCGCTCATGAAATCCAGCATCAACACATCCCAAT 1124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
306 IleAspValIleTyrPheLeuAsnGlu...AsnAlaValGlnLysValAsn 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1125 CAAGCCAGAAATACGACATGTGGAAATCCAGATGGCCACGATATCTCC 1174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
321 eLysProGlnValAspArgTyrArgLeuLysAsnGlyArgArgIleIle 338
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1175 TTTCTTGTGAGAGCGCGCTTCTTAACCTGGCGCTACAGTCAAGTCA 1224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
338 euleuAlaGluGlyArgLeuValAsnLeuGlyCysAlaMetGlyHisPro 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1225 TCTTTCGTTATGTCATGATTCATCAACCAACAGACACTGCTCAAGTCA 1274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIleG 371
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1275 CCTTACGAAAAAGAGAAATCTCGAAGAAAGATTTACACACTTCCGA 1324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```





43	ATTGCCACATCAACCTCCATCTTCTCGGCGGTAGGAACCTAACCTCTC	92
77	VallyasnnlltElysglnAlaIglubheglayaggluilegluileAl	93
93	TGAGAGAGAAATGCGAGGCTTATGGTTCTTCGTAGCGTTATCCGCTT	142
93	agluGlnaspetsbeteralaleullelserleuurygysaraglaInglyG	110
143	CTAAGCCATTGAAGGCTGCAGAAATCTCGTTCCCTCCACATMGACAGTC	192
110	IuIyrsProleAlaIglYAlaIyIsIleValIglYcysThrIstIletrAla	126
193	CAGACAGCGCTCCATGCTGAGACACTCCACAGCTCTTGCTGATGTCAG	242
127	GlnthralaValIleullegluthrleucysalaleuIyAlaIglncysAr	143
243	ATGGGCTTCCTGCACACATCTTCTTACACAGATACAGCCGCTGCTCTA	292
143	gtrPserIaIcysnnlIetYserThrGlnasnGluAlaIAlaIAlaI	160
293	TCGTTGTGCGGCCAACAGACACACACAGAAAGCCGCTATCCAGATC	342
160	eulaA.....GluAlaIyAlaIAlaVal	167
343	TTCCGCTGGAAGGGCGGAACAACCTCCAGATATCTGGGAGAACATACCG	392
168	PheAlatrIpySglYgluSerclInuspaspheTrIptIrcysIleAspar	184
393	CGCTCTCACTAGGCCAGATGTCIACAAGGCCACAGCAGGTTGTGCATGATG	442
184	gCysAlaIasnmet...AspGlyTrpGlnAlaIasnmetIleleuAspaspG	200
443	GTTGGGATGATCACTCCATCCGATCTCCACAGGCGTTCGAATTCGAACAGCC	492
200	IyglYAspIeuthrIstIrtIyValIcysIyIyIstYr.....	211
493	GGTGCTGTTCCAGAGCCAAACAGAGAGCTGACAACTCGAATACCGCTGCGT	542
211	.....	211
543	TCTTGTACACTCAAGCAGCTGTCTCAACCAAGACAGAACCACTGGCACA	592
212	.....ProAsnIvalPheIyIyIyS.....	217





```

?      LENGTH: 138 amino acids
?      TYPE: amino acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      ORIGINAL SOURCE:
?      ORGANISM: TOBACCO SHH PROTEIN
US-08-930-894-5

```

[illegible]

alignment\_block: 7f5-08-0320-804-E  
7f5-08-7559-990-1

Align seg 1/1 to: US-08-930-894-5 from: 1 to: 138

```

340 GTCCTGGCCCTGGAGAGGGCCAAACACTCCACAGATCTAGTGGGAGAAACATA 388
13 ValPhealatrPlygSlYclunthreugIngluTyrTriprcysThrGl 29
390 CCGGCTCTACATGGCCAGATGGTCAAGGCCACAGCAAGTTGTGCATG 439
29 uatrgalaleuasprrPrgIyProglYclYglPProaspLeuIleValAspa 46
440 ATGGTGGTGGAGTCTACAGCTCCATCTCCAGAGGGCTTC..... 477
46 spglYglYAspAlatrhrleuLeuIleHisgluIglYalYalysalaIglu 62
478 GAA'TTTCGAACACCCGGTGGCTGGTTCOCAGACCCAAACGAAGCTGACAACT 527
63 GlutYrAlatIalysSerGlYlysLeuProaspProSerSerThrAspSna 79
528 CGAATPCCGGCGCGCTTCTGCTACACTCAAGACAGGCTTCCAAACCAAGACA 577
79 IGlunPheInleuVal...ThrlleIleArgaspIglYleuLysThrAsp 95
578 AGAACCACTGGCACACAGATTGCTGCCGGCATTAACGGTTCGCCAAGAG 627
95 roleuLysTYrThrglunetLysglunAtrgluValAlglYAlaSerIglu 111
628 ACAACAACAGGTGTCCACCGCTCTACACAGCTCGAAGAAAGAGGCAAACT 677
112 ThrThrThrglYValYsArgleuYrGlmetGlnalaaSnglyhrle 128
678 CCTCTCCACACCAATCAAGTCAACAGAGCT 708
128 uleuPheProAlaIleSnaValAlaAspSer 138

```

seq\_name: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:US-08-930-894-7

```

seq=documentation_block:
? Sequence 7, Application US/08930894
? Patent No. 6037524
?
? GENERAL INFORMATION:
?
? APPLICANT: GREENLAND, Andrew James
?
? APPLICANT: DRAPER, John
?
? APPLICANT: SKIPSEY, Marc
?
? APPLICANT: WARNER, Simon
?
? TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
?
? NUMBER OF SEQUENCES: 10
?
? CORRESPONDENCE ADDRESS:
?
? ADDRESSEE: Pillsbury Madison & Sutro
?
? STREET: 1100 New York Avenue, N.W.
?
? CITY: Washington
?
? STATE: D.C.
?
? COUNTRY: U.S.A.
?
? ZIP: 20005-3918
?
? COMPUTER READABLE FORM:
?
? MEDIUM type: Floppy disk

```

```

?      COMPUTER: IBM PC compatible
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: MS Word
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER: US/08/930,894
?      FILING DATE: 09-Oct-1997
?      CLASSIFICATION: 800
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: PCT/GB96/00882
?      FILING DATE: 10-Apr-1996
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: GB 9507381.3
?      FILING DATE: 10-Apr-1995
?      INFORMATION FOR SEQ ID NO: 7 :
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 139 amino acids
?          type: amino acid
?          STRANDEDNESS: single
?          TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      ORIGINAL SOURCE:
?      ORGANISM: WHEAT SHH PROTEIN (1)
US-08-930-894-7

```

alignment_scores:		
Quality:	307.50	
Ratio:	3.045	
Percent Similarity:	68.707	
	Length:	147
	Gaps:	2
	Percent Identity:	46.259

```
alignment_block: 78-00 030 001 7
```

Align seg 1/1 to: US-08-930-894-7 from: 1 to: 139

```

280 GCGCGTGGTGCATGCGTGGTTCGCGCCCAACAGGCAACAGCAAGAAAGCCAGC 329
      |||||.....:|||||.....:
2  AAAAAAAAAAAAAAAAAAGTspserAlaser..... 12
330 CCGATGCCAGCTTCGCGCTGGAAAGGCGCAACACTGCCAATACTGGG 379
      |||||.....:|||||.....:
13 .....ValPheAlaIrrPLyGLyAluThrLeuIngInGLyTrr 26
380 AGAACACATATACCGCGCCTTCACATATGCGCAGATGGTGAAGGCCACAGCAG 429
      |||.....:||||.....:
26 IrcYstHrGLuArgLAlaLeuAspTrpGLyProGLyLylLeuAspLeu 42
430 GTTGTCAATGATGGTGGTGAATGCTACACTCTCATCTCCAAAGGCGCTTC.. 477
      |||||.....:||||.....:
43 ILeValAspAspGLyLysPrrHrThrLeuLeuLeuHsGLuGLyVal 59
478 .....GAATTCGAAACCGCGCGGCGCTTCACAGGCAACAGAG 517
      |||||.....:|||||.....:
59 sAlaGLuGLuLylurGLyLysThrGLyLysMetProAspProHrSet 76
518 CTGACAACTCGAATATACCGCTGCGCTTCTTGCTACACTCAAGCAGGCTTTC 567
      |||||.....:|||||.....:
76 hrAspAsnAlaGLuPheIngInLlLeValLeuThrLleLeArgAspGLyLeu 92
568 AACCAAGCAGAAAGCACTGGGCACACACATGCTGGCGGCATGACCGTGT 617
      ::|||.....:||||.....:
93 LysValAspProHrLysTyrArgLysMetLysAspArgLlLeValGLy 109
618 TTCGGAAGAGACAACAACAGGTGTCACACGCGCTTCACAGCTCAAGCAAG 667
      |||||.....:|||||.....:
109 lSerGLuGLuThrThrThrGLyValLysArgLeuTyrGLuMetGLuAla 126
668 AGGCGCAACTCTTCCTCCACACCTCAACGTCACAGCGCT 708
126 snAsnSerLeuLeuPheLeuThrLleAsnValAlaAspSer 139

```

seq\_name: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:US-08-204-740-7

## seq\_documentation\_block:

```

; Sequence 7, Application US/08204740
; Patent No. 5753432
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgrett & Wilcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,740
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5753432nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-204-740-7

```

## alignment\_scores:

```

Quality: 287.00 Length: 76
Ratio: 4.100 Gaps: 0
Percent Similarity: 92.105 Percent Identity: 72.368

```

## alignment\_block:

```
US-09-759-990-1 x US-08-204-740-7 ..
```

```
Align seg 1/1 to: US-08-204-740-7 from: 1 to: 97
```

```

37 TACAGAAATGCGACATCAACCTTCATGTCGCGCGTAGAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22 TyrIysValAlaAspIleGlyLeuAlaIatTpgIyArgIysAlaLeuAs 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 CCTTGCTGAGAGGAAGTAATGCGAGCTTATGTTCTTGTGAGCGTTATT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgGluArgTyrS 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 CCGCTTCTAAGCCATGTAAGGCTGTCAGAAATCTGTCGTCCTCCACATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 erAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 71
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 ACAGTCGACAGACGCGCTCATGAGACACTCACAGCTCTTGCTGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 ThrValGluThrAlaValLeuIleGluThrLeuValThrLeuGlyAlaG 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 TGTCAATGGGCTTCCTGCAACATCTTC 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

## 88 uValGlnTrpSerSerCysAsnIlePhe 97

```
seq_name: /cgn2_6/prodata/2/iaa/6A_COMB.pep:US-09-081-167A-7
```

## seq\_documentation\_block:

```

; Sequence 7, Application US/09081167A
; Patent No. 6083745
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,167A
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083745nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-167A-7

```

## alignment\_scores:

```

Quality: 287.00 Length: 76
Ratio: 4.100 Gaps: 0
Percent Similarity: 92.105 Percent Identity: 72.368

```

## alignment\_block:

```
US-09-759-990-1 x US-09-081-167A-7 ..
```

```
Align seg 1/1 to: US-09-081-167A-7 from: 1 to: 97
```

```

37 TACAGAAATGCGACATCAACCTTCATGTCGCGCGTAGAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22 TyrIysValAlaAspIleGlyLeuAlaIatTpgIyArgIysAlaLeuAs 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 CCTTGCTGAGAGGAAGTAATGCGAGCTTATGTTCTTGTGAGCGTTATT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgGluArgTyrS 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 CCGCTTCTAAGCCATGTAAGGCTGTCAGAAATCTGTCGTCCTCCACATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 erAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 71
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 ACAGTCGACAGACGCGCTCATGAGACACTCACAGCTCTTGCTGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

87 CCTGTGAGAAAGAAATGCCAGGCTTATGTTCTTGAGGCTTAT 136
   ::::::::::::::::::::::::::::::::::::::::::::::
38 pIleAlaGluAnsGluMetProGlyLeuMetArgMetArgGluArgTyrS 55
137 CCGCTTCAAGCCATTGAGGGGTGTCAGAAATCTGTGTTCCCTCCACATG 186
   ::::::::::::::::::::::::::::::::::::::::::::::
55 eRAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 71
187 ACAGTCCAGACGCGGTCCTCATCGAGACACTCAGAGCTTGTGGCTGCA 236
   ::::::::::::::::::::::::::::::::::::::::::::::
72 ThrValGlnThrAlaValLeuIleGlnThrLeuValThrLeuGlyAlaG 88
237 TGTCAAGATGGCTTCTCTGCAACATCTTC 264
   ::::::::::::::::::::::::::::::::::::::::::::::
88 uValGlnTrpSerSerCysAsnIlePhe 97

```

seq\_name: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:PCT-US95-02521-7

```

seq_documentation_block:
; Sequence 7, Application PC/TUS9502521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02521
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-02521-7

```

alignment\_scores:

Quality:	287.00	Length:	76
Ratio:	4.100	Gaps:	0
Percent Similarity:	92.105	Percent Identity:	72.368

alignment\_block:

US-09-759-990-1 x PCT-US95-02521-7 ..

Align seg 1/1 to: PCT-US95-02521-7 from: 1 to: 97

```

37 TACAGAAATGCCAGATCAACCTCATGTTCTCGCCGTAAGGAATTAC 86
   :::::::::::::::::::::: ::::::::::::::
22 TyrLysValAlaAspIleGlyLeuAlaIleArgIleGlyAlaLeuAs 38
87 CTTTGTGAGAAAGAAATGCCAGGCTTATGTTCTTCTGAGCGTTATT 136
   :::::::::::::::::::::: ::::::::::::::
38 pIleAlaGluAnsGluMetProGlyLeuMetArgMetArgGluArgTyrS 55
137 CCGCTTCAAGCCATTGAGGGGTGTCAGAAATCTGTGTTCCCTCCACATG 186
   ::::::::::::::::::::::::::::::::::::::
55 eRAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 71
187 ACAGTCCAGACGCGGTCCTCATCGAGACACTCAGAGCTTGTGGCTGCA 236
   ::::::::::::::::::::::::::::::::::::::
72 ThrValGlnThrAlaValLeuIleGlnThrLeuValThrLeuGlyAlaG 88
237 TGTCAAGATGGCTTCTCTGCAACATCTTC 264
   ::::::::::::::::::::::::::::::::::::::
88 uValGlnTrpSerSerCysAsnIlePhe 97

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-204-740-9

seq\_documentation\_block:

; Sequence 9, Application US/08204740

; Patent No. 5753432

; GENERAL INFORMATION:

; APPLICANT: Gudkov, Andrei

; APPLICANT: Kazarov, Alexander

; APPLICANT: Mazo, Ilya

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods for Identifying Genetic

; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant

; TITLE OF INVENTION: Growth in Cancer Cells

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Allgretti & Wilcoff, Ltd.

; STREET: 10 S. Wacker Drive, Suite 3000

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/204,740

; FILING DATE: 04-MAR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5753432nan, Kevin E

; REGISTRATION NUMBER: 35,303

; TELEPHONE: 312-715-1234

; TELEFAX: 312-715-1234

; TELEX: 910-221-5317

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 95 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-204-740-9

alignment\_scores:

Quality:	271.00	Length:	76
Ratio: <td>3.985 <td>Gaps: <td>0</td> </td></td>	3.985 <td>Gaps: <td>0</td> </td>	Gaps: <td>0</td>	0
Percent Similarity: <td>89.474 <td>Percent Identity: <td>69.737</td> </td></td>	89.474 <td>Percent Identity: <td>69.737</td> </td>	Percent Identity: <td>69.737</td>	69.737

alignment\_block:

US-09-759-990-1 x US-08-204-740-9 ..

Align seg 1/1 to: US-08-204-740-9 from: 1 to: 95

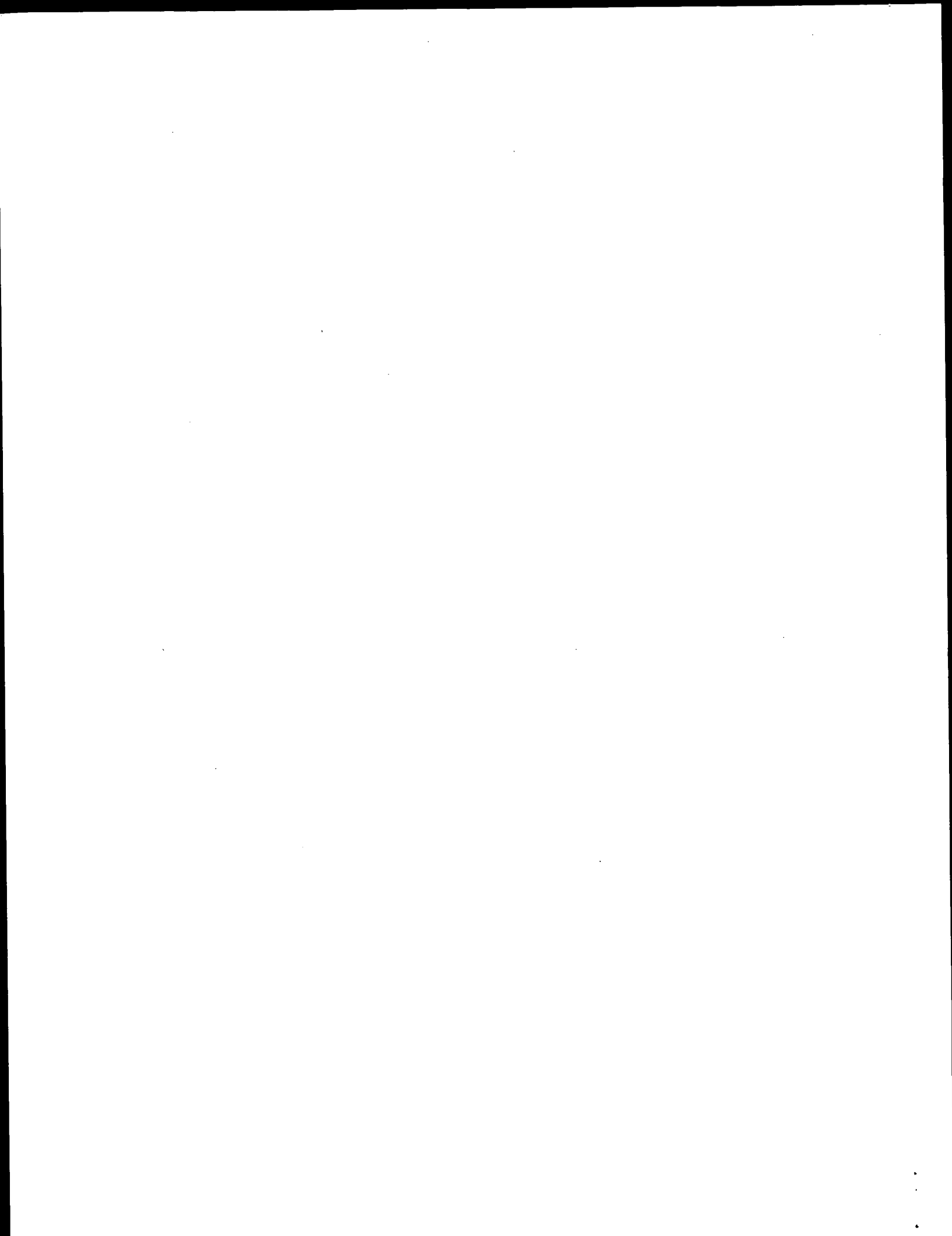
```

37 TACAGAAATGCCAGATCAACCTCATGTTCTCGCCGTAAGGAATTAC 86
   :::::::::::::::::::::: ::::::::::::::
20 TyrLysValAlaAspIleGlyLeuAlaIleArgIleGlyAlaLeuAs 36
87 CTTTGTGAGAAAGAAATGCCAGGCTTATGTTCTTCTGAGCGTTATT 136
   :::::::::::::::::::::: ::::::::::::::
36 pIleAlaGluAnsGluMetProGlyLeuMetArgMetArgGluMetTyrS 53
137 CCGCTTCAAGCCATTGAGGGGTGTCAGAAATCTGTGTTCCCTCCACATG 186
   ::::::::::::::::::::::::::::::::::::::
53 eRAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuArgMet 69
187 ACAGTCCAGACGCGGTCCTCATCGAGACACTCAGAGCTTGTGGCTGCA 236
   ::::::::::::::::::::::::::::::::::::::
70 ThrValGlnThrAlaValLeuIleGlnThrLysValAlaLeuGlyAlaG 86

```



237 TGTTCAGATGGGCTTCCTGCAACATCTTC 264  
:::|||||:::|||||  
86 uAlaArgTrpSerSerCysAsnIlePhe 95



OM of: US-09-759-990-1 to: SwissProt\_40:\* out\_format : pfs

Date: Aug 17, 2002 1:35 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=framer_n2p.model -DEV=slp
-O/cgn2_1/USPTO_spool/US0975990/runat_13082002_153926_28128/app_query.fasta_1.1541
-DB=SwissProt_40 -QPMT=fastan -SUFFIX=n2p.rsp -CAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOPEXT=0.000
-CAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000
-DELOP=6.000 -RGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELEX=7.000 -START=1 -MATRIX=blosum62
-TRANS=human4.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0975990_@CGL1_1_39 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLFP3 -WAIT_THREADS=1
```

Search information block:

```
Query: US-09-759-990-1
Query length: 1461
Database: SwissProt_40:*
Database sequences: 105224
Database length: 3819550
Search time (sec): 49.660000
```

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len
SwissProt_40:SAHH_TRIVA + 2534.00	3020.64	3.1e-161		
SwissProt_40:SAHH_MTCU + 1550.50	1845.36	8.7e-96		
SwissProt_40:SAHH_PETC + 1544.00	1837.77	2.4e-95		
SwissProt_40:SAHH_ARATH + 1537.00	1829.41	6.9e-95		
SwissProt_40:SAHH_TOBAC + 1533.00	1824.63	1.3e-94		
SwissProt_40:SAHH_WHEAT + 1524.00	1813.88	5.0e-94		
SwissProt_40:SAHH_CATRO + 1520.00	1809.10	9.3e-94		
SwissProt_40:SAHH_MESCR + 1518.00	1806.71	1.3e-93		
SwissProt_40:SAHH_PHASS + 1509.00	1795.95	5.0e-93		
SwissProt_40:SAHH_MEDSA + 1499.00	1784.00	2.3e-92		
SwissProt_40:SAHH_RHOCA + 1369.00	1629.14	1.0e-83		
SwissProt_40:SAHH_RHOSH + 1359.00	1617.19	4.8e-83		
SwissProt_40:SAHH_PLAF7 + 1332.00	1584.59	3.0e-81		
SwissProt_40:SAHH_RAT + 1286.00	1530.67	3.4e-78		
SwissProt_40:SAHH_MOUSE + 1278.00	1521.11	1.2e-77		
SwissProt_40:SAHH_SCHPO + 1274.00	1516.29	2.1e-77		
SwissProt_40:SAHH_HUMAN + 1260.00	1499.58	1.8e-76		
SwissProt_40:SAHH_ANOGA + 1252.00	1490.62	5.7e-76		
SwissProt_40:SAHH_DROME + 1251.00	1488.85	7.2e-76		
SwissProt_40:SAH1_XENLA + 1242.50	1478.65	2.7e-75		
SwissProt_40:SAH2_XENLA + 1238.50	1473.87	4.9e-75		
SwissProt_40:SAHH_LEIDO + 1236.00	1470.79	7.2e-75		
SwissProt_40:SAHH_DICDI + 1221.50	1453.63	6.6e-74		
SwissProt_40:SAHH_YEAST + 1217.50	1448.42	1.2e-73		
SwissProt_40:SAHH_CAEEL + 1203.50	1431.96	1.0e-72		
SwissProt_40:SAHH_PNECA + 1142.00	1358.41	1.3e-68		
SwissProt_40:SAH2_DROME + 1081.50	1284.78	1.4e-64		
SwissProt_40:SAHH_SULSO + 1032.00	1225.72	2.6e-61		
SwissProt_40:SAHH_AQUAE + 906.50	1077.56	6.1e-53		
SwissProt_40:SAHH_METJA + 797.00	946.70	1.2e-45		
SwissProt_40:SAHH_SYNY3 + 794.00	943.19	1.9e-45		
SwissProt_40:SAHH_ARCFU + 781.00	927.42	1.4e-44		
SwissProt_40:SAHH_PYRFO + 761.00	904.00	2.9e-43		
SwissProt_40:SAHH_PYRAB + 739.50	877.93	7.9e-42		
SwissProt_40:SAHH_PYRAB + 736.50	874.34	1.3e-41		
SwissProt_40:SAHH_METTH + 735.50	873.12	1.5e-41		
SwissProt_40:SAHH_THEMA + 734.50	872.05	1.7e-41		
SwissProt_40:SAHH_STRFR + 696.50	826.96	5.7e-39		
SwissProt_40:SAH2_YEAST + 372.00	451.24	1.6e-17		
SwissProt_40:MSB2_YEAST + 206.50	229.88	3.2e-06		

SwissProt\_40:AGAL\_YEAST + 202.00 230.33 5.4e-06 725 ! P32323 saccharomyces cere  
SwissProt\_40:YM96\_YEAST + 200.50 224.06 7.7e-06 1140 ! Q04893 saccharomyces cer  
SwissProt\_40:YGLF\_YEAST + 190.50 219.30 2.9e-05 551 ! P53214 saccharomyces cere  
SwissProt\_40:YGLX\_HSVB + 178.50 201.31 0.0002 757 ! P28968 equine herpesvirus  
SwissProt\_40:YHC8\_YEAST + 177.50 202.84 0.0002 605 ! P38739 saccharomyces cere

seq\_name: SwissProt\_40:SAHH\_TRIVA

seq\_documentation\_block:

```
ID SAHH_TRIVA STANDARD; PRT; 486 AA.
AC P51540;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
hydrolyase) (AdoHcyase).
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalidea; Trichomonadida; Trichomonas.
OX NCBI_TaxID=5722;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WAA38;
RX MEDLINE=97047381; PubMed=88923301;
RA Bagnara A.S., Tucker V.E., Minotto L., Howes E.R., Ko G.A., /
RA Edwards M.R., Dawes I.W.;
RT "Molecular characterisation of adenosylhomocysteinase from
Trichomonas vaginalis."
RL Mol. Biochem. Parasitol. 81:1-12(1996).
CC -!- FUNCTION: ADENOSYLMHOCYSTEINE IS A COMPETITIVE INHIBITOR OF
S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
THEREFORE ADENOSYLMHOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CONCENTRATION OF ADENOSYLMHOCYSTEINE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
+ L-homocysteine.
CC -!- COFACTOR: NAD (BY SIMILARITY).
CC -!- PATHWAY: ACTIVATED METHYL CYCLE.
CC -!- SIMILARITY: BELONGS TO THE ADENOSYLMHOCYSTEINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U40872; RAC47319.1; --
CC HSSP; P10760; IB3R.
CC InterPro; IPR000043; Ado_hcyase.
CC Pfam; PF00670; AdoHcyase; 1.
CC PROSITE; PS00738; AdoHcyase_1; 1.
CC PROSITE; PS00739; AdoHcyase_2; 1.
CC Hydrolase; NAD; One-carbon metabolism.
CC NP_BIND; 267 298 NAD (POTENTIAL).
CC SEQUENCE 486 AA; 53439 MW; 1403G6750587427C CRC64;

```

alignment\_scores:

Quality: 2534.00 Length: 486  
Ratio: 5.225 Gaps: 0  
Percent Similarity: 99.794 Percent Identity: 98.971

alignment\_block:

US-09-759-990-1 x SAHH\_TRIVA ..  
Align seg 1/1 to: SAHH\_TRIVA from: 1 to: 486  
1 ATGGCTTGCAATACCTACTGGTCTCCATTCGAGTACAGAAATGCCGA 50  
|||||  
1 MetAlaCysLysSerProAlaGlyAlaProPheGluTyArgIleAla 17  
51 CATCAACCTCCATGTTCTCGCCCGCTAGGAACTTACCTTGCTGAGAAG 100

|||||  
17 pIleAsnLeuHisValLeuGlyArgLysGluLeuThrLeuAlaGluLysG 34  
101 AAATGCCAGGTCTTATGGTTCTTCGTCAGCGGTTATTCGGCTTCTTAAGCCA 150  
34 luMetProGlyLeuMetValLeuArgGluArgTyrSerAlaSerLysPro 50  
151 TTGAAGGGTGTCAAGATCTCTGGTTCCTCCACATGACAGTCCGACAGC 200  
51 LeuLysGlyValArgIleSerGlySerLeuHisMetThrValGlnThrAl 67  
201 CGTCCCTCATGACACACTCACAGTCTTGGTTCGTGATGTCAGATGGGCTT 250  
67 aValLeuIleGluThrLeuThrAlaLeuGlyAlaAspValArgTrpAlaAs 84  
251 CMTGCAACATCTTCTACACAAGATACAGCCGCTGCTGCTATCCTGTC 300  
84 erCysAsnIlePheSerThrGlnAspThrAlaAlaAlaIleValVal 100  
301 GGCCCAACAGCCACACAGAGAGCCAGCCGGTATCCAGTCTTCGCCCTG 350  
101 GlyProThrGlyThrProGluLysProAlaGlyIleProValPheAlaTr 117  
351 GAAGGGGGAACACTCCAGAAATACCTGGGAGAACACATACCGCGCTCTCA 400  
117 pLysGlyGluThrLeuProGluTyrTrpGluAsnThrTyrArgAlaLeuT 134  
401 CATGCCAGATGGTCAAGGCCACACAGAGTTGTGATGATGATGGTGTAT 450  
134 hrTrpProAspGlyGlnGlyProGlnGlnValValAspAspGlyGlyAsp 150  
451 GCTACACTCTTCATCTCCAAAGGCTTCGAATTCGAAACAGCCGGTCTGT 500  
151 AlaThrLeuLeuIleSerLysGlyPheGluPheGluThrAlaGlyAlaVal 167  
501 TCCAGAGCAACAGAGCTGCAACACTCGAATACCGCTCGCTTCTGCTA 550  
167 lProGluProThrGluAlaAspAsnLeuGluTyrArgCysValLeuAlaTr 184  
551 CACTCAAGCAGGTCTTCAACCAAGCAACAAACACTGGCCACACAGTTGCT 600  
184 hrLeuLysGlnValPheAsnGlnAspLysAsnHisTrpHisThrValAla 200  
601 CGCGCATGACGGTGTTCGAAAGTCCGAAAGACAAACAGGTGTCCACCGCCT 650  
201 AlaGlyMetAsnGlyValSerGluGluThrThrThrGlyValHisArgLe 217  
651 CTACAGCTCGAAGAGGAGGGCAAACTCTCTTCCAGCCATCAACGCTCA 700  
217 uTyrGlnLeuGluLysGluGlyLysLeuLeuPheProAlaIleAsnValAla 234  
701 AGCAGCGTGTTCAAAGTCCAAAGTTCGATAAACATCTACGGCTGCGGCCAC 750  
234 snAspAlaValThrLysSerLysPheAspAsnIleTyrGlyCysArgHis 250  
751 TCCCTTATCGATGGTATCAACCGTGTCTCGATGTCATGATCGCGGCCAA 800  
251 SerLeuIleAspGlyIleAsnArgAlaSerAspValMetIleGlyGly 267  
801 GACAGCTCTCGTATCGGTTACGGCATGTCGCAAGGCTTCGCTCAAT 850  
267 sThrAlaLeuValIleGlyTyrGlyAspValGlyLysGlyCysAlaGlns 284  
851 CCTCGGTGCCAAGCGCTCGGCTTATCATCATCAAGACATCGACCCCAATC 900  
284 erLeuArgGlyGlnGlyAlaArgValIleIleThrGluValAspProIle 300  
901 TCGGCTCTCCAGCTGCCATCGAAGGTACAGTCCGCGCCGATCGAGGA 950  
301 CysAlaLeuGlnAlaValMetGluGlyTyrGlnValArgIleGluG 317  
951 AGTCGTCAAGGATGTCGATATCTTCGTTACATGACAGGAACATCGCGATA 1000  
|||||

317 uValValLysAspValAspIlePheValThrCysThrGlyAsnCysAspI 334  
1001 TCATCTCTGTGACATGATGCCCCAGATGAAGATAAGGCTATTGTGCGT 1050  
334 leIleSerValAspMetMetAlaGlnMetLysAspLysAlaIleValGly 350  
1051 AACATCGCCCACTTCGATPACGAAATTGATACAGATGGCCTCATGAATA 1100  
351 AsnIleGlyHisPheAspAsnGluIleAspThrAspGlyLeuMetLys 367  
1101 CCCAGGATCAAGCACATCCCAATCAAGCCAGACATACGATGTGGGAAT 1150  
367 rProGlyIleLysHisIleProIleLysProGluTyrAspMetIleP 384  
1151 TCCAGATGGCCACGCTATCCTCTCTTCTGCTGAGGGCGGCTTCTTAAC 1200  
384 heProAspGlyHisAlaIleLeuLeuAlaGluGlyArgLeuLeuAsn 400  
1201 CTTGGCTCGCTACAGTCAACCCATCTTTCGTTATGTCATGTCATTCAC 1250  
401 LeuGlyCysAlaThrGlyHisProSerPheValMetSerMetSerPheTh 417  
1251 AAACAGACACTCGCTCAGCTCGACCTCTACGAAAGAGAGAGAAATCTCG 1300  
417 rAsnGlnThrLeuAlaGlnLeuAspLeuTyrGluLysArgGlyAsnLeuG 434  
1301 AGAGAAGGTTTACACACTTCGGAAGCATCTCGATGAAGAGTGGCTCGC 1350  
434 luMetLysValTyrThrLeuProLysHisLeuAspGluGluValValArg 450  
1351 CTCACCTCGGATCTCGATGTCACCTTACAAAGCTTACACAGCTTACACA 1400  
451 LeuHisLeuGlySerLeuAspValHisLeuThrLysLeuThrGlnLysG 467  
1401 GCGTACTACATCAAGTTCACCTGAGGCTGCTTACAAAGTCTGATGCTT 1450  
467 nAlaAspTyrIleAsnValProValGluGlyProTyrLysSerAspAla 484  
1451 ACCGTTAT 1458  
484 yrArgTyr 486  
seq\_name: SwissProt\_40:SAHH\_MYCTU  
seq\_documentation\_block:  
ID SAHH\_MYCTU STANDARD; PRT; 494 AA.  
AC O08364;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine  
hydrolase) (AdoHcyase).  
GN AHCY OR SAHH OR RV3248C OR MT3346 OR MTCY20B11.23C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RA MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence.";  
RL Nature 393:537-544(1998).  
RN [2]

334 ATCCAGTCTTCGCTGGAGGCGCAACACTCCCAAGATACTGGGAA 383  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
119 ValProValPheAlaTrpLysGlyGluThrLeuGluGluTyrTrpAl 135  
::: :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
384 CACATACCGCGCTTCACATGGCCAGCATGCTCAAGGCCA...CAGCAGG 430  
::: :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
135 aAlaGlnMetLeuThrTrpAspProAspLysProAlaAsnMetI 152  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
431 TTGTGCGATGGTGGTGATGTACACTCCTCATCTCTCAAGGGCTTCGAA 480  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
152 leLeuAspAspGlyGlyAspAlaThrMetLeuValLeuArgGlyMetGln 168  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
481 TTCGAAACAGCGGCTCTTCAGAGCCAAACAGAGCTGACAACTTCGA 530  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
169 TyrGluLysAlaGlyValValProAlaGluGluAspAspProAlaG 185  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
531 ATACCGCTCGCTTCTGTCTACACTCAAGCAGGCTTCAACCAACAGCA 580  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
185 utTrpLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLys 202  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
581 ACCACTGGCACACAGTTCCTGCCGGCATGAACGGTGTTCCTCGAAGACA 630  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
202 spLysTrpThrLysIleAlaGluSerValLysGlyValThrGluGluThr 218  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
631 ACAACAGCTGTCCACCGCTCTACAGCTTCGAGAGGAGGCAACCTCCT 680  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
219 ThrThrGlyValLeuArgLeuTyrGlnPheAlaAlaAlaGlyAspLeuAl 235  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
681 CTTCGAGCCATCAACGTCACAGCGCTGTACAAAGTCCAAGTTCGATA 730  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
235 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 252  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
731 ACATCTACGGCTGCCGCACCTCCCTATCGATGGTATCAACGCTGCTTC 780  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
252 snLysTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 268  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
781 GATGTCATGATCGCGCGCAAGACAGCTCTCGCTCATGGTTACGGCGATGT 830  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
269 AspAlaLeuIleGlyGlyLysLysValLeuIleCysGlyTyrGlyAspVa 285  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
831 CGSCAAGGGTGGCTCAATCCCTCCGTCGCAAGCGCTCGGGTTATCA 880  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
285 lGlyLysGlycysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 302  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
881 TCACAGAACTCGACCAATCTCGCTCTCCAGGCTGCCATGGAAGCTAC 930  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
302 alThrGluIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 318  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
931 CAGGTCCGGCGCATCGAAGAGTCGTCAAGGATGTCGATATCTTCGTTAC 980  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
319 AspValValThrValGluGluAlaIleGlyAspAlaAspIleValValTh 335  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
981 ATCCACAGCAAACTGGCATATCATCTCTGTGTACATGATGGCCAGATGA 1030  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
335 rAlaThrGlyAsnLysAspIleIleMetLeuGluHisIleLysAlaMetL 352  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
1031 AGGATAAGGCTATTGTGGTAACATCGGCCACTTCGATAACGAAATTCAT 1080  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
352 ysAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 368  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
1081 ACAGATGGCTCATGAAATACCCAGGCATCAAGCACATCCCAATCAAGCC 1130  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
369 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysPr 384  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
1131 AGAATACGACATGTGGGAATTCCAGAT...GGCACGCTATCCCTCCTTC 1177  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
384 oGlnValAspLeuTrpThrPheGlyAspThrGlyArgSerIleIleValL 401  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
1178 TTGCTGAGGGCGCCTCTTAACCTTGGCTGCGGTACAGGTACCCCATCT 1227  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
401 euSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 417  
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

----- This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sb.ch](mailto:license@sib-sb.ch)).

EMBL; M81885; AAA33856.1; -

```

198 ArgLeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTy 214
654 CCAGCTCGAAGAGGAGGCAAACTCCTCTCCAGGCATCAACGTCACAG 703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
214 rGlnMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnA 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
704 ACGCTGTACAAAGTCCAAAGTTCGATACATCATACGGTCCGCCACTCC 753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
231 spSerValThrLysSerLysPheAspAsnLeuTyrgLysArgHisSer 247
754 CTTATCATCGTATCAACCGCTCTCCGATGTCATGATCGGGCGGCAAGAC 803
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
248 LeuProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVa 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
804 AGCTCTCGTCATGGGTTACGGCGATGTCCGCAAGGCTCGCTCAATCCC 853
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
264 lAlaLeuIleAlaGlyTyrgLysValGlyLysGlyCysAlaAlaAlaAl 281
854 TCGGTGCCAAGCGCTCGCTTATCATCATACAGAACTCGACCAATCTGC 903
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
281 etLysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCys 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
904 GCTCTCCAGGTCGCATGGAGGCTACCAGTCCCGCGCATCGAGGAAGT 953
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
298 AlaLeuGlnAlaThrMetGluGlyLeuGlnValLeuProLeuGluAspVa 314
954 CPTCAAGATGTCATATCTTCGTTACATGCACAGAACTCGCATATCA 1003
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
314 lValSerGluValAspIlePheValThrThrThrGlyAsnLysAspIleI 331
1004 TCTCTGTGTGATGATGCCAGATGAAGGATGAAGCTATTCTCGGTAC 1053
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
331 leMetValSerAspMetArgLysMetLysAsnAlaIleValCysAsn 347
1054 ATGGCGCATCTCGATACGAATTTATACATGATGCGCTCATGAATACCC 1103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
348 lIleGlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyPr 364
1104 AGGCATCAACACATCCCATCAAGCCAGATACGACATGCGGAATGCC 1153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
364 oGlyValLysArgIleThrIleLysProGlnThrAspArgTrpValPheP 381
1154 CAGAT...GCCACGCTATCCTCTCTTCTGTGAGGCGCGCTCTTAA 1200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
381 roAspThrGlyArgGlyIleIleIleLeuAlaGluGlyArgLeuMetAsn 397
1201 CTGTGCTGCTACAGTCAACCATCTCTTCTGTTATGTCAATGTCAATCAC 1250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
398 LeuGlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheTh 414
1251 AAACACGACACTCGCTCAGTCAGCTCAGCTCTAGGAAAGAGA..... 1294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
414 rAsnGlnValIleAlaGlnLeuGluLeuIleThrPasnGluLysSerGlyL 431
1295 ATCTCGAAGAGAGGTTTACACACTTCGAGCATCTCGATGAAGAGTC 1344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
431 ystYrGlyLysLysValTyValLeuProLysHisLeuAspGluLysVal 447
1345 GCTCGCTCCACTCGATCTCTCGATCTCCACTTACAAAGCTTACACA 1394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
448 AlaAlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuSerLy 464
1395 GAAGCAGGCTGACTACATCAACCTTCAGTTAGGGTTCCTTACAAAGTCTG 1444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
464 sAspGlnAlaAspTyIleSerValProValGluGlyProTyTrLysProA 481
1445 ATGCTTACGCTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
481 lAhIstYrArgTytr 485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq\_name: SwissProt\_40:SAHH\_ARATH

seq\_documentation\_block:

```

ID AC 023255; 081847; STANDARD; PRT; 485 AA.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHCyase).
GN SAHH OR AT4G13940 OR DL3010W.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERECTA;
RA Belbahri L., Elleuch H., Villarreal R., Inze D., Thomas D.,
RA Thomasset B.;
RT "The isolation of an Arabidopsis thaliana cDNA clone encoding S-
RT adenosyl-L-homocysteine hydrolase";
RL (in) Plant Gene Register PGR99-139.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=98121113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,
RA Wedler H., Wedler E., Wambutt R., Weitzenecker T., Pohl T.M.,
RA Terry N., Giesen J., Villarreal R., de Clerck R., van Montagu M.,
RA Lechamy A., Auborg S., Gy I., Kreis M., Lao N., Kavanagh T.,
RA Hempel S., Kottler P., Entian K.-D., Rieger M., Schaeffer M., Funk B.,
RA Mueller-Auer S., Silvey M., James R., Montfort A., Pons A.,
RA Puigdomenech P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P.,
RA Piravandi E., Obermaier B., Hilbert H., Dueterhoeft A., Moores T.,
RA Jones J.D.G., Eneva T., Palme K., Benes V., Rechman S., Ansoorge W.,
RA Cooke R., Berger C., Delsen M., Voet M., Volckaert G., Mewes H.-W.,
RA Klosterman S., Schueller C., Chaiwatzis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT Arabidopsis thaliana";
RL Nature 391:485-488(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Dueterhoeft A., Stiekema W., Entian K.-D., Terry N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenecker T., Bothe G., Ransperger U., Hilbert H., Braun M.,
RA Holzer B., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooljman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Bernier S., Hempel S., Feldpausch M., Lamberth S., van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarreal R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Partmann B., Grandrath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lechamy A., Aubourg S.,
RA Chedfor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis C.,

```

RA Frishman D., Haase D., Lemcke K., Meves H.-W., Stocker S.,  
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedhia N., Gnoj L., Schütz K., Huang E., Spiegel L.,  
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Kramer J., Fulten D., Mardis E., Dente M., Pepin K., Hillier L.,  
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antoniou B., Zidanik M., Strong C., Sun H., Lamar B., Jordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
RA Granat S., Shohdy N., Hasedawa A., Hameed A., Lodhi M., Johnson A.,  
RA Chen E., Marra M., Martienssen R., McCombie W.R.,  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RT thaliana.";  
RT Nature 402:769-777(1999).  
[4]  
RN SEQUENCE OF 19-485 FROM N.A.  
RC STRAIN-CV LANDSBERG ERECTA;  
RC Zhang H., Forde B.G.;  
RT "Identification of novel nitrate-inducible genes from Arabidopsis.";  
RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF  
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;  
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE  
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR  
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine  
CC + L-homocysteine.  
CC -!- COFACTOR: NAD.  
CC -!- PATHWAY: ACTIVATED METHYL CYCLE.  
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; Z97335; CAB10173.1; -;  
CC EMBL; AF059581; AAC14714.1; -;  
CC EMBL; AL161537; CAB78436.1; -;  
CC EMBL; Z97059; CAB09795.1; -;  
CC HSSP; P10760; 1B3R.  
CC InterPro; IPR000043; Ado\_hcyase.  
CC Pfam; PF00670; AdoHcyase; 1.  
CC DR PROSITE; PS00738; ADHCHYASE\_1; 1.  
CC DR PROSITE; PS00739; ADHCHYASE\_2; 1.  
CC KW Hydrolase; NAD; One-carbon metabolism.  
CC FT NP\_BIND 263 294 NAD (POTENTIAL).  
CC FT CONFLICT 80 80 E -> Q (IN REF. 4).  
CC FT CONFLICT 96 96 A -> R (IN REF. 4).  
CC FT CONFLICT 392 392 E -> Q (IN REF. 4).  
CC FT CONFLICT 460 460 T -> R (IN REF. 4).  
CC SQ SEQUENCE 485 AA; 53378 MW; 1113270A0F46C86C CRC64;

## alignment\_scores:

Quality: 1537.00 Length: 487  
Ratio: 3.901 Gaps: 4  
Percent Similarity: 80.903 Percent Identity: 62.012

## alignment\_block:

US-09-759-990-1 x SAHH\_ARATH ..

Align seg 1/1 to: SAHH\_ARATH from: 1 to: 485

19 ACTGGTGCTCCATTCAGGTACAGAAATTCGCGACATCAACCTCCATGTTCT 68

8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24  
69 CGCGCGTAAGGAACCTTACCCCTTGCTGAGCAAGAAATCCAGCTCTTATGG 118  
24 eGlyArgGluLeuLeuGluAlaGluValGluMetProGlyLeuMetA 41  
119 TTCTTCGTGAGCGTTATTCCCTCTTAAGCCNATTAAGAGGTGTCAGATC 168  
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile 57  
169 TCTGGTTCCTCCACATCAGTCCAGACGCGCTCTCATCGAGACACT 218  
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74  
219 CACAGCTCTTGGTCTGCTGATGTGATGGCTTCCGCAACATCTTCTCTA 268  
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSert 91  
269 CACAAGATACAGCGCTGCTGCTATCTGTTGTCGGCCCAACAGGCACCA 318  
91 hrGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105  
319 GAGAGCCAGCGGTATCCAGTCTTGGCTGGAAGGGCGAAACACTCCC 368  
106 .....ValPheAlaTrpLysGlyGluThrLeuG1 115  
369 AGAATCTGGGAGAACATACCCGCTCTCATATGGCCAGATGGTCAAG 418  
115 nGluTyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 132  
419 GCCCACAGCAGGTGCTGATGTGTTGATGCTACACTCTCATCTCTCC 468  
132 lypTroAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148  
469 AAGGCTTCGAA.....TTCGAAACAGCGCGTCTGTTCCAGA 506  
149 GluGlyValLysAlaGluLulPheGluLysThrGlyGlnValProAs 165  
507 GCCAACAGAGCTGACACCTCGAATACCGCTGCTGCTGCTGCTGCTGCTG 556  
165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIleL 182  
557 AGCAGGTCTTCAACCAAGAAAGAACTGGCACACAGTGTGCTGCGCGC 606  
182 yGluGlyLeuGlnValAspProLysLysThrHisLysMetLysGluArg 198  
607 ATGACAGGTGTTTCCGAGAGACAAACAGAGGTGTCCACCGCTCTTACCA 656  
199 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuYrGl 215  
657 GCTCAGAGAGGAGGCAAACTCTCTTCCAGCCATCAACGTCACAGCAGC 706  
215 nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAspS 232  
707 CTGTTACAAAGTCCAGTTCGATACATCTACGCTGCGCCCTCCCTT 756  
232 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248  
757 ATCGATGTATCAACCGGTCTTCCGATGTGATGTCGCGCGCAGACAGC 806  
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265  
807 TCTCGTCATGGGTACGCGCATGTCGCAAGGGCTGCGCTCAATCCCTCC 856  
265 aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetL 282  
857 GTGGCCAAAGCGCTCGCGTTATCATCACAGAACTGACCCCAATCTGCGCT 906  
282 yThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298  
907 CTCAGGCTGCCATGGAAGGCTACCAGGTCCGCGCATCGAGGAAGTCGT 956



```

299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValva 315
957 CAAGATGTCGATATCTTGGTTCATGACGAGGAACTGCGATATCATCT 1006
315 lserGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
1007 CTGTGTCATGATGCGCCAGATGAGGATAAGGCTATTGTCGGTACATC 1056
332 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
1057 GCCACTTCGATAACGAATTCATGACAGATGGCTCATGAAATACCCAGG 1106
349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrProG 365
1107 CATCAAGCAGATCCCAATCAAGCCAGATACGATGTCGGATATCCAG 1156
365 yValLysArgIleThrIleLysProGlnThrAspArgTyrPvalPheProG 382
1157 ATGCCACAGCT...ATCCCTCTTCTGCTGAGGCGGCTTCTTACCTT 1203
382 luthrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
1204 GCTGCGCTACAGTCAACCCATCTTCTGTTATGTCAATGTCATTACAAA 1253
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAATC 1297
415 nGlnValIleAlaGlnLeuGluLeuIleTyrAsnGluLysAlaSerGlyLysT 432
1298 TCGAGAGAAGAGTTTACACACTTCGAGACATCTCGATCAAGAAGTCGT 1347
432 yrGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
1348 CGCTCCACCTCGATCTCGATGTCACCTTACAAAGCTTACAGAA 1397
449 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAs 465
1398 GCAGGCTGACTACATCAAGCTTCAGTTCAGGTCCTTACAGTCTGATG 1447
465 pGlnSerAspTyrValSerIleProIleGluGlyProTyrLysProProH 482
1448 CTTACGCTTAT 1458
482 istYrArgTyr 485

seq_name: SwissProt_40:SAHH_TOBAC
seq_documentation_block:
ID SAHH_TOBAC STANDARD; PRT; 485 AA.
AC P50248;
AT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase) (cytokinin binding protein CBP57).
GN SAHH.
OS Nicotiana tabacum (Common tobacco), and
OS Nicotiana sylvestris (Wood tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097, 4096;
RN [1]
RN SEQUENCE FROM N.A.
RC SPECIES=N.tabacum; STRAIN=CV, BRIGHT YELLOW 4;
RA Tanaka H., Masuta C., Kataoka J., Kuwata S., Koizumi A., Noma M.;
RT "Inducible expression by plant hormones of S-adenosyl-homocysteine
RT hydrolase gene from Nicotiana tabacum during early flower bud
RL formation in vitro."
RL Plant Sci. 113:167-174(1996).
RN [2]
RN SEQUENCE FROM N.A.

```

```

RC SPECIES=N.sylvestris;
RA Mitsui S., Wakasugi T., Sugiura M.;
RT "A cDNA encoding the 57 kDa subunit of a cytokinin-binding protein
RT complex from tobacco: the subunit has high homology to S-adenosyl-L-
RT homocysteine hydrolase."
RL Plant Cell Physiol. 34:1089-1096(1993).
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D45204; BAA08142.1; -;
CC DR EMBL; D49804; BAA23164.1; -;
CC DR EMBL; D16138; BAA03709.1; -;
CC DR HSSP; P10760; 1B3R.
CC DR InterPro; IPR000043; Ado_hcyase.
CC DR Pfam; PF00670; AdoHcyase; 1.
CC DR PROSITE; PS00738; ADOHCYASE_1; 1.
CC DR PROSITE; PS00739; ADOHCYASE_2; 1.
CC KW Hydrolase; NAD; One-carbon metabolism.
CC FT NP_BIND; 263 294 NAD (POTENTIAL).
CC SQ SEQUENCE 485 AA; 53104 MW; AA6D6844E9DFOA5C CRC64;

```

alignment\_scores:

Quality: 1533.00 Length: 482  
Ratio: 3.951 Gaps: 4  
Percent Similarity: 80.498 Percent Identity: 63.278

alignment\_block:

US-09-759-990-1 x SAHH\_TOBAC ..

Align seg 1/1 to: SAHH\_TOBAC from: 1 to: 485

```

34 GAGTACAGATTCGCGACATCAACCTCCATGTTCTCGCGCGTAAGGAAC 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GluTyrLysValLysAspMetSerGlnAlaAspPheGlyArgLeuGlu 29
84 TACCTTGTCTGAGAGAAATGCCAGGCTTATGTTCTTCGTGAGCGTT 133
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgThrGlu 46
134 ATTCCGCTTCTTAAGCAATGAAGGCTGTCAAGATCTCTGTTCCCTCC 183
:||||:||||:||||:||||:||||:||||:||||:||||:||||:
46 heGlyProSerGlnProPheLysGlyAlaLysIleThrGlySerLeuHis 62
184 ATGACAGTCCAGACAGCGCTCTCATCGAGACACTCACAGCTCTTGGTGC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 79
234 TGATGTGATGGCTTCTCTCAACATCTTCTTACACAAGATACAGCGC 283
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 aGluValArgTyrCysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATCGTTGTCGCGCCCAACAGCAGCACACAGAGAGCCAGCG 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 laAlaAlaIleAlaArgAspSerAlaAla..... 105
334 ATCCAGTCTTCGCTGGAAGGCGGAACACTCCAGATATCTGGAGAA 383

```



alignment\_block:

US-09-759-990-1 x SAHH\_WHEAT ..

Align seg 1/1 to: SAHH\_WHEAT from: 1 to: 485

```

19 ACTGGTCTCCATTCGAGTACAGAAATTCGCGACATCAACCTCCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerGlyArgGluTyrLysValLysAspLeuPheGlnAlaAspPh 24
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGGCCGTAAGAACTTACCTTCGCTGAGAAGAAATGCCAGGCTTATGG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMet 41
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTCTCTGACGCTTATTCGCTTCTTAAGCCATTGAAGGTGTGCAATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGGTTCCCTCCACATCACAGTCCAGACAGCGCTCTCATCGAGACACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTTGGTCTGATGTCAGATGGGCTCTCCGCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTyrCysSerCysAsnIlePheSerS 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAAGATACAGCGCTGCTGCTATCTGTTGTCGGCCCAACAGGCACACA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 erGlnAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319 GAGAAGCAGCGGCTATCCCACTCTTCGCTGGAGGGGCAACACTCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaIlePheGlyGluThrLeuG 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATCTGGAGACACATACCCGCTCTCACATGGCCAGATGTCACAG 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 uGluTyrTrpCysThrGluArgCysLeuAspTyrGlyValGlyG 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 GCCCAGCAGCAGGTGTGATGATGGTGTGATGTCACACTCTCTCATCTCC 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 LyrProAspLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHis 148
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 AGGGGCTC.....GAATTCGAACACGCGGTGTCTCCAGA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 GluGlyValLysAlaGluGluPheGluLysSerGlyLysValProAs 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
507 GCCAACAGAGCTGACACCTCGAATACCGCTGCTGCTTCTGCTACATCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 ProGluSerThrAspAsnProGluPheLysIleValLeuThrIleAla 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
557 AGCAGGTCTTCAACCAAGACAAGAACCTCGGCACACAGATGCTGCGCGC 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 rgAspGlyLeuLysThrAspAlaSerLysTyrArgLysMetLysGluArg 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
607 ATGACAGGTGTTCCGAGAGACAACAAGGTGTCACCGCCCTTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 LeuValGlyValSerGluThrThrThrGlyValLysArgLeuTyrG 215
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 GCTCAGAGAGAGGGCAACTCTCTTCCAGCCATCAAGTCAACAGC 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 nMetGlnGluSerGlyThrLeuLeuPheProAlaIleAsnValAsnAs 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707 CTGTTACAAAGTCCAAAGTTCGATACATCTACCGCTGCGCCCACTCCCT 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757 ATCGATGCTATCAACCGCTCTCCGATGTCATGATCGCGGCAAGACAGC 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVal 265
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807 TCTGTCATGGTTACGGGATGTGCGGCAAGGCTCGCTCAATCCCTCC 856
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 aValValCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaLeuLeu 282

```

```

857 GTGCCCAAGCGCTCGCTGTTATCATCAAGAACTCGACCCCAATCTCGCT 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 ysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
907 CTCAGGCTGCCATGGAAGCTACAGGTCCGCGCATCGAGGAAGTCTGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 LeuGlnAlaLeuMetGluGlyIleGlnIleLeuThrLeuGluAspValVa 315
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
957 CAAGGATGTCGATATCTCTTACATCACAGAAACTGCGATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 lSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleLeu 332
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1007 CNGTTCACATGATGGCCCATCAAGCATAGGCTATTGTGCGTAACATC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1057 GCCCACTTCGATAACGAAATGTATACAGATGGCTCATGAAATACCCAG 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 GlyHisPheAspAsnGluIleAspMetAsnGlyLeuGluThrTyrPro 365
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1107 CATCAGCACATCCCAATCAAGCCAGAAATACGACATGTGGAAATCCAG 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 yValLysArgIleThrIleLysProGlnThrAspArgTyrValPhePro 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1157 ATGGCCAC...GCTATCCCTCTCTGCTGAGGCGCCCTCTTAACTT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 luthrLysThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1204 GGCTCGCTACAGGTACACCATCTTTCGTTATGTCATGTCATTCAAAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThr 415
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1254 CCAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAAATC 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 nGlnValIleAlaGlnLeuGluLeuTrpAsnGluLysAlaSerGlyLys 432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1298 TCGAGAAGAAGGTTTACACACTTCGAGCATCTCGATGAAGAAGTCTGT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 yrGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1348 CGCTCCACCTCGGATCTCGATGTCACCTTACAAAGCTTACACAGAA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449 AlaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLys 465
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1398 GCAGGCTGACTACATCAAGTTCAGGTGAGGTCTTACAGTCTGATG 1447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465 rGlnSerAspTyrIleSerIleProIleGluGlyProTyrLysLeuArg 482
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1448 CTTACCGTTAT 1458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 eutyArgTyr 485
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq\_name: SwissProt\_40:SAHH\_CATRO

seq\_documentation\_block:

ID SAHH\_CATRO STANDARD; PRT; 485 AA.

AC P35007;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine

DE hydrolase) (AdoHcyase).

GN SAHH.

OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asceridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;

OC Vincetoxicaceae; Catharanthus.

OX NCBI\_TaxID=4058;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94218405; PubMed=8165255;

RA Schroeder G., Waitz A., Hotze M., Schroeder J.;  
RT "cDNA for S-adenosyl-L-homocysteine hydrolase from Catharanthus  
EL roseus";  
PL Plant Physiol. 104:1099-1100(1994).  
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF  
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE  
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR  
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE.  
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine  
CC + L-homocysteine.  
CC -1- COFACTOR: NAD.  
CC -1- SUBUNIT: ACTIVATED METHYL CYCLE.  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- INDUCTION: BY STRESS.  
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; 226881; CAA81527.1; -;  
DR PIR; S38379; S38379.  
DR HSP; P10760; 1B3R.  
DR InterPro; IPR000043; Ado\_hcysase.  
DR InterPro; IPR002025; NAD\_binding.  
DR Pfam; PF00670; AdoHcysase; 1.  
DR PROSITE; PS00738; ADHCYASE\_1; 1.  
DR PROSITE; PS00739; ADHCYASE\_2; 1.  
DR Hydrolase; NAD; One-carbon metabolism.  
KW NP\_BIND; 263 294 NAD (POTENTIAL).  
FT SEQUENCE 485 AA; 53233 MW; 50CC0E99A9F66C51 CRC64;  
SQ

alignment\_scores:  
Quality: 1520.00 Length: 488  
Ratio: 3.848 Gaps: 6  
Percent Similarity: 80.943 Percent Identity: 62.295

alignment\_block:  
US-09-759-990-1 x SAHLH\_CATRO ..

Align seg 1/1 to: SAHLH\_CATRO from: 1 to: 485

```

19 ACTGGTGTCTCCATTCCAGTACAGAAATTCGCGACATCAACCTCCATGTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24
69 CGGCGTAAAGAACTTACCTTCCTCAGAGGAAATCCAGGCTTATGG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuMets 41
119 TTCTCGTGGAGGCTTATTCGGTCTTAAGCCATTGAAGGCTGCAGAAATC 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 erCysArgAlaGluPheGlyProSerGlnProPheLysGlyAlaLysile 57
169 TCTGGTTCCTCCACATGACAGTCCAGACAGCGCTCCTCATCGAGACACT 218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
219 CACAGCTCTGTGGTGCATGATCGAGTGGCTTCTCGCAACATCTTCTCTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSert 91
269 CACAAGATACAGCGCTGCTGCTATCTGTTGTCGGCCCAACAGGCACACA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 hrGlnGluHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319 GAGAACCCACCGGTATCCCGATCTTCGGCTCGGAAGGCGCAACACTCCC 368

```

```

106 .....ValPheAlaTrpLysGlyGluThrLeuG1 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATACATGGAGACACATACCGCGCTCTCACATGG....CCAGATGGTC 415
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 nGluTyrTrpCysThrGluArgAlaLeuAspTrpGlyProAspGly. 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 AAGCCACACAGCAGGTTCTCGATGATGGTGGTATGCTACATCCCTCATC 465
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 ..GlyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeu 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466 TCCAAAGGGCTTC.....GAATTCGAAACAGCCGGCTGCTTCC 503
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 HisGluGlyValLysAlaGluGluTyrLysLysAsnGlyAlaLeuPr 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 AGAGCCACACAGAGCTGACAACCTCGAATACCGCTGCTTCTTCTGTACAC 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 oAspProSerThrAspAsnAlaGluPheGlnIleValLeuThrIle 181
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
554 TCAAGCAGGTTCTTCAACCAAGACAGAACCACTGGCACACAGTTGCTGCC 603
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 leArgAspGlyLeuLysSerAspProThrLysTyrThrArgMetLysGlu 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
604 GGCATGAACGGTGTTCCTCCAGAGACAAACACAGCTGCCACCGCTCTA 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 ArgLeuValGlyValSerGluGluThrThrGlyValLysArgLeuTy 214
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
654 CCAGCTCGAGAAGGAGGCAAACTCCTCTCCAGCGCATCAACGTCAACG 703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 rGlnMetGlnAlaAsnGlyThrLeuLeuPheProAlaIleAsnValAsn 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
704 AGCGTGTATCAAGTCCAAAGTTCGATAACATCTACGGCTGCCGCCACTCC 753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 spSerValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSer 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
754 CTTATCGATGGTATCAACCGTGTCTCGATGTCATGATCGCGCGCAAGAC 803
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
248 LeuProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVa 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
804 AGCTCTCGTCATGGTTACGGCGATGTGGCAAGGGCTGCGCTCAATCCC 853
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 lAlaValValAlaGlyTyrGlyAspValGlyLysGlyCysAlaAlaL 281
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
854 TCCGTGGCCAAAGCGCTCGGTATCATCATCAGAACTCGACCCCAATCTGC 903
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 euLysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCys 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
904 GCTCTCCAGGCTGCCATGGAAGCTACAGGTCCCGCGCATCGAGGAAGT 953
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298 AlaLeuGlnAlaThrMetGluGlyLeuGlnValLeuThrLeuGluAspVa 314
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
954 CGTCAAGGATGTCGATATCTCGTTACATGCACAGGAACCTCGGATATCA 1003
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 lValSerGluAlaAspIlePheValThrThrGlyAsnLysAspIle 331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1004 TCTCTGTGTGACATGATCGCCAGATGAAGGATAAGGCTATTGTCGGTAAC 1053
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331 leMetValAspHisMetArgLysMetLysAsnAlaIleValCysAsn 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1054 ATCGGCCACTTCGATAACGAAATTCATACATGCGCTCATGAATACCC 1103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
348 lLeGlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrPr 364
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1104 AGCATCAAGCACATCCCAATCAAGCCAGAAATACGACATGTGGGAATCC 1153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 oGlyValLysArgIleThrIleLysProGlnThrAspArgTrpValPheP 381
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1154 CAGATGCCCAAGCT...ATCCTCTCTTCTGTCGAGGCGCGCTTCTTAAC 1200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
381 roAspThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsn 397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201 CTGCGCTCGGCTACAGGTCACCGCTCTTCTGTTATGTCATATGTCATTCAC 1250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

398  LeuGlyCysAlaThrGlyHisProSerPheValMetSerCysSerPhe 414
1251  AAACCAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAA 1294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414  rAsnGlnValIleAlaGlnLeuGluLeuTprAsnGluArgLysThrGlyL 431
1295  ATCTCGAGAAGAAGTTTACACACTTCGGAACCATCTCGATGAAGAAGTC 1344
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431  yStrYGluLysValTyValLeuProLysHisLeuAspGluLysVal 447
1345  GCTCGCTCCACCTCGGATCTCTCGATCTCCACCTTACAAAGCTTACACA 1394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448  AlaAlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThryl 464
1395  GAACGAGCTGACTACATCAACGTTCCAGTTGAGGGTCCTTACAGTCTG 1444
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
464  sAspGlnAlaAspTyrIleSerValProIleGluGlyProTyrLysProA 481
1445  ATGCTTACCGTTAT 1458
|||||:
481  laHisTyrArgTyr 485

```

seq\_name: SwissProt\_40:SAHH MESC

seq\_documentation\_block:

ID	SAHH_MESCR	STANDARD;	PRT;	485 AA.
10	50205			

P93253; AC  
15-DEC-1998 (Rel. 37, Created)  
15-DEC-1998 (Rel. 37, Last sequence update)  
15-DEC-1998 (Rel. 37, Last annotation update)  
Adenosylhomocysteinease (EC 3.3.1.1) (S-adenosyl-L-homocysteine hydrolase) (AdoHcyase).  
SAHH.  
Mesembryanthemum crystallinum (Common ice plant).  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Caryophyllaceae; Alizacoaceae; Mesembryanthemum.  
NCBI\_TaxID=3344;  
CC

RP SEQUENCE FROM N.A.

Michalowski C.B., Bohnert H.J.;

submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.

-1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine + L-homocysteine

-!- COFACTOR: NAD.

- I - PATHWAY: ACTIVATED METHYL CYCLE.

- /- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
- /- SIMILARITY: BY ANALOGY.

----- SIMILARITY: BELONGS TO THE ADENOSYLMOMOCYSTEINASE FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----

-----  
EMPT. 1170766. 1170766.

EMBL; U/9766; AAB38499.1; -.  
HSSP: P10760. 1B3P

InterPro; IPR000043; Ado hcyase

Pfam; PF00670; AdoHcyase; 1.

PROSITE; PS00738; ADONCYASE\_1; 1.

PROSITE; PS00739; ADOHCYASE\_2; 1.  
Hydrolase: NAD: One-carbon meta-

NP_BIND	263	294	NAD	Metabo
NP_BIND	263	294	NAD	Metabo

Q	SEQUENCE	485 AA;	53178 MW;	2C

100

\_\_\_\_\_

```

alignment_scores:
  Quality: 1518.00      Length: 487
  Ratio: 3.824          Gaps: 4
  Percent Similarity: 81.520  Percent Identity: 61.602

alignment_block:
US-09-759-990-1 x SAHH_MESCR ..

Align seg 1/1 to: SAHH_MESCR from: 1 to: 485

19  ACTGTTGTCCTCCATTTCAGTATACAGAAATTCGCGACATCAACCTCCATGTTCT 68
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
8  ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
69  CGGCGGTAAAGNACTTACCCTTGCTGAGAGGAAATGCCAGGTCTTATGG 118
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
24  eGlyArgLeuGluLeuGluAlaGluValGluMetProGlyLeuMeta 41
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
119 TTTCTTCGTGAGCGTTATTCGCTTCTTAAGCCATTGAAGGTGTCAGATC 168
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
41  laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaLysile 57
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
169 TCTGTTTCCCTCCATGACAGTCCAGACAGCGCTCTCATCGAGACACT 218
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
58  ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
219 CACAGCTCTTGCTGATGTCAGATGGCTTCCTGCAACATCTTCCTA 268
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
74  uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
269 CACAAGATACAGCCGCTGCTGCTATCTGTTGTGCGCCCAACAGCAGCACCA 318
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
91  hrGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
319 GAGAAGCCAGCGGTATCCCACTTTCGCTGGAAGGCGGAACACTCCC 368
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
106 .....ValPheAlaTrpLysGlyGluThrLeuGI 115
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
369 AGAATACTGGGAACACATACCCGCTCTCATGTGCCAGATGGTCAAG 418
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
115 nGluTyrTrpCysThrGluArgAlaLeuAspTrpGlyAlaGlyGlyG 132
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
419 GCCACAGCAGGTTGTCGATGATGGTGTGATGTCACACTCCTCATCTCC 468
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
132 lProAspLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHis 148
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
469 AAGGGGTTCC.....GAATCGAAACAGCGGTGCTGTCCACA 506
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
149 GluGlyValLysAlaGluGluGluTyrGluLysAsnGlyThrIleProAs 165
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
507 GCCACAGAGCTGACACCTCGNATACCGCTGCTTCTTGCTACACTCA 556
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
165 proThrSerThrAspAsnProGluPheGlnLeuValLeuGlyLeuIleA 182
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
557 AGCAGGTCTTCAACCAAGACAAGAACCACTGGCACACAGTTCCTCCGCG 606
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
182 rgAspSerLeuLysValAspProLysArgTyrHisLysMetLysThrArg 198
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
607 ATGAACGGTGTTCGAAGACACACACAGGTGTCCACCGCTCTACCA 656
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
199 LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTyrGI 215
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
657 GCTCAGAGAGGAGGCAAACTCCTCTTCCAGCCATCAAGCTCAACGACG 706
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
215 nMetGlnAlaThrGlyThrLeuLeuPheProAlaIleAsnValAsnAspS 232
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
707 CTGTTTACAAATCCAAAGTTCGATAACATCTACGGCTGCCGCGCACTCC 756
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
232 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
757 ATCATGGTATCAACCGCTCTCCGATGTCATCATCGGCGGACAGACAGC 806
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValGI 265
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

807 TCTGTCATGGTACGAGATCGCGAAGGCGCTCAATCCCTCC 856
      :::::|||||
265 yvalValCysGlyTyrGlyAspValGlyGlyCysAlaLeuAlaLeuL 282
      :::::|||||
857 GTGGCAAGGCGCTCGGTTATCATCAGAACTCGACCAATCTCGCT 906
      :::::|||||
282 ysAlaAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
      :::::|||||
907 CTCAGGCTGCATCGAAGCTACCAAGCTACCGCGCATCGAGAACTCGT 956
      :::::|||||
299 LeuGlnAlaLeuMetGluGlyPheGlnIleLeuThrLeuGluAspValVa 315
      :::::|||||
957 CAAGGATGCGATATCTGTTACATGCACAGAAACTCGGATATCATCT 1006
      :::::|||||
315 lserGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
      :::::|||||
1007 CTGTTGACATGATGGCCAGATGAAGTAAAGCTATTGCGGTAACATC 1056
      :::::|||||
332 etValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 348
      :::::|||||
1057 GGCCACTTCGATAACGAATTTGATACAGATGGCTCATGAAATACCCAGG 1106
      :::::|||||
349 GlyHisPheAspGlnIleAspMetLeuGlyLeuGluAsnTyrProG 365
      :::::|||||
1107 CATCAAGCATCCCAATCAAGCCAGATACGATGCTGGGAATTCACG 1156
      :::::|||||
365 yvalLysArgIleThrIleLysProGlnThrAspArgPheValPheProG 382
      :::::|||||
1157 ATGGCCAC...GCTATCTCTCTCTGCTGAGGCGCGCTTCTTAACCTT 1203
      :::::|||||
382 luthrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetLysLeu 398
      :::::|||||
1204 GGCTGCGCTACAGTCAACCTCTTTGTTATGTCATGTCATTCACAAA 1253
      :::::|||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
      :::::|||||
1254 CCAGACATCGCTCAGCTCGACCTCTACGAAAGAGA...GGAAATC 1297
      :::::|||||
415 nglInValIleAlaGlnLeuGluLeuTrpAsnGluArgAlaSerGlyLysT 432
      :::::|||||
1298 TCGAAGAAGGTTTACACTTCGGAAGCATCTCGATGAGAAGTCTGCT 1347
      :::::|||||
432 yrGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
      :::::|||||
1348 CGCTCCACCTCGATCTCGATCTCCACCTTACAAAGCTTACACAGAA 1397
      :::::|||||
449 AlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuSerLysAs 465
      :::::|||||
1398 GCAGGCTGACTACATCAAGCTTCCAGTTGAGGCTTACAGAGTCTGATG 1447
      :::::|||||
465 pGlnAlaAspTyrIleSerValProValGluGlyProTyrLysProAlaH 482
      :::::|||||
1448 CTTACCGTTAT 1458
      :::::|||||
482 lstyArgTyr 485
      :::::|||||

seq_name: SwissProt_40:SAHL_PHASS
seq_documentation_block:
ID: SAHL_PHASS STANDARD; PRT: 485 AA.
AC P50249; 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
CN SAHL.
OS Phalaenopsis sp.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
OC Epidendroideae; higher Epidendroideae; Vandaeae; Aeridinae;
OC PhalaenopsisLs.
```

```

NCBI_TaxID=36900;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95177653; PubMed=7872785;
RA Preisig-Mueller R., Gnau P., Kindl H.;
RT "The inducible 9, 10-dihydrophenanthrene pathway: characterization
RT and expression of bilenzyl synthase and S-adenosylhomocysteine
RT hydrolase.";
RL Arch. Biochem. Biophys. 317:201-207(1995).
CC -I- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -I- COFACTOR: NAD (BY SIMILARITY).
CC -I- PATHWAY: ACTIVATED METHYL CYCLE.
CC -I- INDUCTION: BY INFECTION WITH B. CINEREA.
CC -I- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X79905; CAA56278.1; -.
CC HSP: P10760; 1B3R.
CC InterPro: IPR000043; Ado_hcyase.
CC Pfam: PF00670; AdoHcyase; 1.
CC PROSITE: PS00738; ADHCHYASE_1; 1.
CC PROSITE: PS00739; ADHCHYASE_2; 1.
CC Hydrolase; NAD; One-carbon metabolism.
CC NP_BIND 263 294 NAD (POTENTIAL).
CC SEQUENCE 485 AA; 53141 MW; 1EB2CA5AD63AF233 CRC64;

alignment_scores:
Quality: 1509.00 Length: 482
Ratio: 3.889 Gaps: 4
Percent Similarity: 80.498 Percent Identity: 62.656

alignment_block:
US-09-759-990-1 x SAHL_PHASS
Align seg 1/1 to: SAHL_PHASS from: 1 to: 485

34 GAGTACAGAATTGCCGACATCAACCTCCATGTTCTCGGCCCTAAGGAAC 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
13 GluTyrLysValLysAspLeuSerGlnAlaAspPheGlyArgLeuGluI 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84 TACCTTGCTGAGAAGAAATGCCAGGCTTATGTTCTCTCGTAGCGTT 133
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgAlaGlu 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134 ATTCCTGCTTCTAAGCCATTGAAGGTTGTGAGAAATCTCTGGTTCCCTCCAC 183
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
46 heGlyProSerGlnProPheLysGlyAlaArgIleSerGlySerLeuHis 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184 ATGACAGTCCAGACAGCGCTCTCATCGAGACACTCACAGCTCTTGGTGC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 TGATGTCAGATGGGCTCTCTGCAACATCTTCTCTACACAAGATACAGCG 283
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
79 aGluValArgTyrCysSerCysAsnIlePheSerThrGlnAspHisAla 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
284 CTGCTGCTATCGTTCTCGGCCCAACAGCCACACAGGAGAGCCGCCGCT 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
96 laAlaAlaIleAlaArgAspSerAlaAla..... 105
```

334 ATCCAGTCTTCGCGCTGGAGGCGGAAACACACTCCCAAGAACTACTGGGAGAA 383  
106 .....ValPheAlaTrpIysGlyGluThrLeuGlnGluTrpTrpCy 120  
384 CACATACCGCGCTTCACATGCCAGATGTCAGGCGCCACAGCAGGTTG 433  
120 sThrGluArgCysLeuGluTrpGlyAlaGlyGlyProAspLeuIleV 137  
434 TCGATGATGGTGGTATGCTACACTCCTCTCATCTCCCAAGGCGCTTC..... 477  
137 alAspAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 153  
478 .....GAATTCGAACAGCGGTCTGCTGCTTACCTCAAGAGGCTTC..... 477  
154 GluGluGluTrpGluLysAsnGlyLysIleProAspProAlaSerThrAs 170  
522 CAACCTCGAATACCGCTCGCTTCTGCTACACTCAAGCAGGCTCTCAACC 571  
170 pAsnAlaGluPheGlnIleValLeuGlyLeuIleArgAspSerLeuSerV 187  
572 AAGACAAGAACCACTGGCACACAGTGTCTGCCGGGATGAACGGTGTTC 621  
187 alAspProLysLysTyrArgMetLysGluArgLeuValGlyValSer 203  
622 GAAGACAGACAACAGGTGTCCACGCGCTCTACCACTGCGAGAGGAGG 671  
204 GluGluThrThrGlyValLysArgLeuGlyGlnMetGlnTrpSerG 220  
672 CAACCTCCTCTCCAGCCATCAACGTCAAGCAGCGCTGTACAAAGTCCA 721  
220 yThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSerL 237  
722 AGTTCGATAACACTAGCGTCCGCGCGCACTCCCTTATCGATGATCAAC 771  
237 ySphAspAsnLeuLysTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253  
772 CGTGTCCGATCATGATGATCGCGCGGCAAGACAGCTCTCGTCATGGGTTA 821  
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValValCysGlyTy 270  
822 CGCGATGTCGCGCAAGGTCGCGCTCAATCCCTCGTGGCGCAAGCGCTC 871  
270 rGlyAspValGlyLeuGlyCysAlaAlaLeuLysThrAlaGlyAlaA 287  
872 GCGTATCATCAGAACTCGACCAATCTCGCTCTCCAGCTCCCATG 921  
287 rGValIleValThrGluIleAspProIleCysAlaLeuGlnAlaLeuMet 303  
922 GAAGCTACAGTCCGCGCGCATCGAGGAAGTCGTCAAGGATGTCGATAT 971  
304 GluGlyLeuProValLeuArgLeuGluAspValValSerGluAlaAspI 320  
972 CTTCGTTACATGACAGGAACCTGCGATATCATCTCTGTTGACATGATGG 1021  
320 ePheValThrThrGlyAsnLysAspIleIleMetValAspHisMetA 337  
1022 CCAGATGAAGTAAAGGCTATGTCGGTAACATCGGCCACTTCGATAAAC 1071  
337 rGlyMetLysAsnAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353  
1072 GAAATGTATACAGTGGCTCATGAATACCCAGGATCAAGCACATCCC 1121  
354 GluIleAspMetLeuGlyLeuGluSerPheProGlyValLysArgIleTh 370  
1122 AATCAAGCAGAAATACGACATGTGGGAATTCACAGATGGCCAGCT...A 1168  
370 rIleLysProGlnThrAspArgArgValPheProAspThrAsnSerGlyI 387  
1169 TCCTCCTTCTGCTGAGGCGCCGCTTCTTAACCTTCGGCTCGCTACAGGT 1218  
387 leLeuValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403

1219 CACCCATCTTTCGTTATGCTCAATGCTATTCACAAACACAGACACTCGCTCA 1268  
404 HisProSerPheValMetSerSerPheThrAsnGlnValIleAlaG 420  
1269 GCTCGACCTCTAGCAAAAGAGA.....GGAAATCTCCAGAGAAGGTTT 1312  
420 nLeuGluLeuTrpLysGluArgAlaSerGlyLysTyrGluLysValT 437  
1313 ACACATTCGGAAGCATCTCGATGAAGAAGTCGCTCGCTCCACCTCGGA 1362  
437 yValLeuProLysHisLeuAspGluLysValAlaAlaLeuHisLeuGly 453  
1363 TCTCTCGATGTCACCTTACAAAGCTTACACAGAGGAGGCTGACTACAT 1412  
454 LysLeuGlyAlaLysLeuThrLysLeuThrProSerGlnAlaAspTyrI 470  
1413 CAACGTTCCAGTTGAGGTCCTTCAAGTCTGATGCTTACCGTTAT 1458  
470 eSerValProValGluGlyProTyrLysProAlaHisTyrArgTyr 485  
seq\_name: SwissProt\_40:SAHH\_MEDSA  
seq\_documentation\_block:  
ID SAHH\_MEDSA STANDARD; PRT; 485 AA.  
AC P50246;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine  
DE hydrolase) (AdoHcyase).  
GN SAHH.  
OS Medicago sativa (Alfalfa).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
OX NCBI\_TaxID=3879;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. SIRIVER; TISSUE=Stem;  
RX MEDLINE=95201245; PubMed=7894016;  
RA Abrahams S., Hayes C.M., Watson J.M.;  
RT "Expression patterns of three genes in the stem of lucerne (Medicago  
sativa).";  
RL Plant Mol. Biol. 27:513-528(1995).  
CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF  
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;  
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE  
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR  
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine  
CC + L-homocysteine.  
CC -!- COFACTOR: NAD (BY SIMILARITY).  
CC -!- PATHWAY: ACTIVATED METHYL CYCLE.  
CC -!- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L36119; AAB41814.1; -.  
DR HSSP; P10760; 1B3R.  
DR InterPro; IPR000043; Ado\_hcyase.  
DR Pfam; PF00670; AdoHcyase\_1.  
DR PROSITE; PS00738; ADHOCYASE\_1; 1.  
DR PROSITE; PS00739; ADHOCYASE\_2; 1.  
KW Hydrolase; NAD; One-carbon metabolism.  
FT NP\_BIND 263 294 NAD (POTENTIAL).  
FT SEQUENCE 485 AA; 53151 MW; 13EAFDBD527DF59F CRC64;

alignment\_scores:  
 Quality: 1499.00 Length: 482  
 Ratio: 3.853 Gaps: 4  
 Percent Similarity: 80.705 Percent Identity: 61.411

alignment\_block:  
 US-09-759-990-1 x SAHH\_MEDSA ..

Align seg 1/1 to: SAHH\_MEDSA from: 1 to: 485

```

34 GAGTACAGAAATCCCGACATCAACCTCCATGTTCTCGCGCGTAAGGAAC 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GlutyllysVallylsAspMetSerGlnAlaAspPheGlyArgLeuGluI 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 TACCCTTGCTGAGAAGAAATGCCAGGTCTTATGGTCTTCGTCGAGCGTT 133
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 eGluLeuAlaGluValGluMetProGlyLeuMetSerCysArgThrGluP 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 ATTCCGCTTCTAAGCATTTGAAGGTCTCAAGATCTCTGGTTCCTCCAC 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 heglyProSerGlnProPheylsGlyAlaArgIleThrGlySerLeuHis 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 ATGACAGTCCAGACAGCGCTCTCATCGAGACACTCACAGCTCTTGGTGC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 TGATGTCAGATGGCTTCCTGCAACATCTTCTTACACAAGATACAGCGG 283
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 aGluValArgTrpCysSerCysAsnIlePheSerThrGlnAspHisAlaA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 CTGCTGCTATCTGTTGTCGCGCCCAACAGGACACACCAGAGACGCGGT 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 laAlaAlaIleAlaArgAspSerAlaAla..... 105
334 ATCCAGCTCTTCGCTGGAAGGGGAAACACTCCAGATACACTGGAGAA 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTrpLysGlyGluThrLeuGlnGluTrpTrpCy 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 CACATACCGCTCTCACATGCCAGATGGTCAAGGCCACACAGAGTTG 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 sSerGluArgAlaLeuAspTrpGlyProGlyGlyProAspLeuIleV 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434 TCGATGATGGTGGTATGATCTACCTCATCTCCAAAGGGCTTCGAA... 480
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 alAspAspGlyGlyAspValThrLeuLeuIleHisGluGlyVallyAla 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 .....TTCGAACAGCGGTGCTGTTCCAGAGCCCAACAGAGCTGA 521
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 GluGluValPheGluLysThrGlyGlnLeuProAspProSerSerThrAs 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
522 CAACCTCGAATACCGCTCGCTTCTTGTACACTCAAGCGGTCTCAACC 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 pAsnAlaGluMetGlnIleValLeuThrIleLeuArgAspGlyLeuLys 187
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
572 AAGACAAGAACACTGGCACAGTGTGCTGCGCGCATGACGGTGTTC 621
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 hrAspProLysArgTrpGlnLysMetLysThrArgIleValGlyValSer 203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
622 GAAGACACAAACAGGTGCTCCACCGCTCTTACAGCTCGAGAGGAGGG 671
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 GluGluThrThrThrGlyVallylsArgLeuTyrglnMetGlnAlaSerGl 220
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
672 CAACCTCTCTTCCAGCGCATCAACGCTCAACGACGCTGTTCACAAAGTCCA 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 yThrLeuLeuPheProAlaIleAsnValAspSerValThrLysSerL 237
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
722 AGTTCGATACATCTACGCTGCTCCGCTCTTATCGATGGTATCAAC 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 ysPheAspAsnLeuTyrglyCysArgHisSerLeuProAspGlyLeuMet 253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
772 CGTGTCTCGCATGTCATGTCGCGCGCAAGACAGCTCTCGTCATGGGTTA 821
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValValCysGlyTy 270
822 CGCGGATGTGCGCAAGGGCTCGCTCAATCCCTCGTGGCGCAAGCGCTC 871
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 rGlyAspValGlyLysGlyCysAlaAlaAlaLeuLysGlnGlyAlaAla 287
872 GCCTTATCATCACAGAACTGACCCAAATCTCGCTCTCCAGGCTGCCATG 921
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 rgValIleValThrGluIleAspProIleCysAlaLeuGlnAlaLeuMet 303
922 GAAGGCTACAGGTCGCGCGCATCGAGGAAGTCGTCGAAGGATGCGATAT 971
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 GluGlyLeuGlnValLeuThrIleGluAspValIleSerGluAlaAspIle 320
972 CTTCGTTACATGCACAGAAACTGCGATATCATCTCTGTTCACATGATCG 1021
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ePheValThrThrThrGlyAsnLysAspIleIleMetValSerAspMetL 337
1022 CCCAGATGAAGGATAAGGCTATTGTCGCTAACATCGCCACTTCGATAAC 1071
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 yLysMetLysAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353
1072 GAAATTGATACAGATGGCTCATGAATACCCAGGCATCAAGCACATCCC 1121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 GluLeaspMetHisGlyLeuGluThrTyProGlyValLysArgIleTh 370
1122 AATCAAGCCAGAAATACGACATGTCGGAATCCAGATGGCCAGCT...A 1168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 rIleLysProGlnThrAspArgTrpValPheProGluThrLysSerGlyI 387
1169 TCCCTCTTCTGCTGAGGCGCCCTTCTTAACCTTGGCTGCGCTACAGGT 1218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 leIleValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
1219 CACCATCTTCTGTTATGTCATGTCATTCACAAACCAGACACTCGCTCA 1268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaGl 420
1269 GCTCGACCTCTACGAAAGAGA.....GGAAATCTCGAGAAGAGTTT 1312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 nIleGluLeuTrpLysGluLysThrSerGlyLysTyThrGluLysValT 437
1313 ACACACTTCCGAGCATCTCGATGAAGAAGTCGCTCGCTCCACCTCGGA 1362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
437 yrValLeuProLysHisLeuAspGluLysValAlaAlaLeuHisLeuGly 453
1363 TCTCTCGATGCTCCACTTACAAAGCTTACACAGAGCAGGCTGACTACAT 1412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 GlnLeuGlyAlaLysLeuThrLysLeuSerLysAspGlnAlaAspTyrI 470
1413 CAACGTTCCAGTTGAGGTCCTTACAAAGTCTGATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 eSerValProValGluGlyProTyLysProAlaHisTyArgTy 485
seq_name: SwissProt_40:SAHH_RHOCA
seq_documentation_block:
ID SAHH_RHOCA STANDARD; PRT; 463 AA.
AC P28183;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN AHCY.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB1003 / ST LOUIS;
RX MEDLINE=92335291; PubMed=1631127;

```



86 laAlaIaIleAlaAlaIaGly.....Gly 93  
334 ATCCAGTCTTCGCTCGAAGGCGCAACACTCCAGAAATACTGGAGAA 383  
94 ThrProValPheAlaValLysGlyGluThrLeuGluGluTyrTrpAlaTyr 110  
384 CACATACCCGCTCTCACATGGCCAGATGGTCAAGGCCACAGCAGAGTTG 433  
110 rThrAspLysIlePheGlnPheProGluGlyThr...CysAsnMetIleL 126  
434 TCGATCATGGTGGTGATGTCTACTCTCTCATCTCCATCCAAAGGCTTCGAAATTC 483  
126 euAspGlyGlyAspAlaThrLeuTyrIleLeuLeuGlyAlaArgVal 142  
484 GAAACA.....GCCGGTGTGTTCAGAGCCCAACAGAACTGCAACACT 527  
143 GluAlaGlyGluThrAspLeuIleAlaThrProThrSerGluAspGluVal 159  
528 CGAATACCGCTCGGTTCTTGCTACACTCAAGCAGGTCTTCAACCAAGACA 577  
159 I.....Cys.....LeuPheAsnGlnIleL 166  
578 AGAACCA.....TGCAACACA...GTTGCTGCGGCG 606  
166 yLysArgMetValGluSerProGlyTyrPheThrGlnGlnArgAlaAla 182  
607 ATGAACGGTGTTCGGAAGAGACAACAACAGGTCTCCACGGCTCTACCA 656  
183 IleLysGlyValSerGluGluThrThrGlyValHisArgLeuTyrAs 199  
657 GCTCGAGAAGCAGGCAACTCTCTCTCCAGCCATCAACGTCAACGACG 706  
199 pLeuHisLysGlyLeuLeuProPheProAlaIleAsnValAsnAsp 216  
707 CTGTTACAAAGTCCAAGTTCGATAACATCTACGGCTGCGCCACTCCCTT 756  
216 erValThrLysSerLysPheAspAsnLysTyrGlyLysLysGluSerLeu 232  
757 ATCGATGGTATCAACGGTGTTCGATGTCATGTATGATCGCGCGCAACAGAC 806  
233 ValAspGlyIleArgArgAlaThrAspValMetMetAlaGlyLysValAl 249  
807 TCTCGTTCATGGTTACGCGGATGTCGCAAGGCTGCGCTCATCCCTCC 856  
249 aValValCysGlyTyrGlyAspValGlyLysGlySerAlaAlaSerLeuA 266  
857 GTGCGCAAGCGCTCGGTTATCATCACAGAACTCGACCCATCTGCGCT 906  
266 rgGlyAlaGlyAlaArgValLysValThrGluValAspProIleCysAla 282  
907 CTCAGGCTGCCATGGAAGCTACCAGTCCGCGCGCATCTGAGGAAGTCGT 956  
283 LeuGlnAlaAlaMetAspGlyPheGluValValLeuGluAspValVal 299  
957 CAAGGATGTCGATATCTTCGTTACATGCACAGAAACTGCGGATATCATCT 1006  
299 lAlaAspAlaAspIlePheIleThrThrGlyAsnLysAspValIleA 316  
1007 CNGTTACATGATGGCCCGCAGATGAAGATTAAGCTATTGTGCGTAACATC 1056  
316 rgIleGluHisMetArgGluMetLysAspMetAlaIleValGlyAsnIle 332  
1057 GCCCACTTCGATTAACGAATTTGATACAGATGGCTCATGAAATACCCAGG 1106  
333 GlyHisPheAspAsnGluIleGlnValAlaAlaLeuLysAsnHis..... 347  
1107 CATCAAGCACATCCCAATCAAGCAGAAATACGACATGTGGGAATTCACG 1156  
348 ....LysThrThrAsnIleLysAspGlnValAspMetIleGluMetPro 363  
1157 ATGGCCAGCTATCTCTCTTCCTTCGTAGAGGGCGCTCTTAACTTGGC 1206  
363 erGlyAlaArgIleIleLeuLeuSerGluGlyArgLeuLeuAsnLeuGly 379

233 ValAspGlyIleArgAlaThrAspValMetMetAlaGlyLysValAl 249  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

807 TCTCGTCATGGTTTACGGCATGCGCAAGGCCTGCCTCAATCCCTCC 856  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

249 aValValCysGlyTyrGlyAspValGlyLysGlySerAlaLeuA 266  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

857 GTGGCCAAGCGCTCGCGTTATCATCACAGAACCTCGACCCTCTCGCGT 906  
|||||||:::|||||:::|||||:::|||||:::|||||:::|||||

266 rgGlyAlaGlyAlaArgValLysValThrGluValAspProIleCysAla 282  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

907 CTCACGGCTGCCATGGAAAGCTACCAAGTCCGCCGCATCGAGGAAGTCGT 956  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

283 LeuGlnAlaAlaMetAspGlyPheGluValValValLeuGluAspValVa 299  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

957 CRAAGGATGTCGATATCTTCTTGATGCACAGGAACAATCGCATATCATCT 1006  
|||||||:::|||||:::|||||:::|||||:::|||||:::|||||

299 lAlaAspAlaAspIlePheIleThrThrGlyAsnLysAspValIleA 316  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

1007 CTGTTTGACATGATGGCCCCACATCAAGGATAAGCTATTGTGCGTAACATC 1056  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

316 rgIleGluHisMetArgGluMetLysAspMetAlaIleValGlyAsnIle 332  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

1057 GGCACCTTCGATAACGAATTGATACAGATGGCGCTCATGAATACCCAGG 1106  
|||||||:::|||||:::|||||:::|||||:::|||||:::|||||

333 GlyHISPhaAspAsnGluIleGlnValAlaLeuLysAsnHis..... 347  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

1107 CATCAAGCACATCCCACATCAAGCCAGAATACGACATGTGGGAATTCCTCAG 1156  
|||||||:::|||||:::|||||:::|||||:::|||||:::|||||

348 ...LysrTrpThrAsnIleLysaspGlnValAspMetIleGluMetProS 363  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

1157 ATGGCCACGCTATCTCTCTCTCTGTGAGGGCGCGCTCTTTAACCTTGGC 1206  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

363 erGlyAlaArgIleIleIleLeuSerGluGlyArgLeuLeuAsnLeuGly 379  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

1207 TGGCTACAGGTCACCCATCTTTGTTATGTCATTCATTCACAAACCA 1256  
|||||  
380 AsnAlaThrGlyHisProSerPheValMetSerAlaSerPheThrAsnG1 396  
1257 GACACTCGCTCAGCTGACCTCTACGAAAGAGAGAAATCTCGAGAAGA 1306  
1:|||||  
396 nValLeuAlaGlnIleGluLeuThrPheThrLysGlyAlaGluTyrGlnProg 413  
1307 AGCTTTACACATTCGAGACATCTCGATGACGAGAGCTCGCTCGCTCCAC 1356  
|||||  
413 lyValTyrIleLeuProLysSerLeuAspIleValAlaArgLeuHis 429  
1357 CTCGGATCTCTGATCTCCACCTTACAAAGCTTACACAGAGAGGCTGA 1406  
1:|||||  
430 LeuLysIleGlyValLysLeuThrThrLeuArgProAspGlnAlaG1 446  
1407 CTACATCAAGCTTCAGGCTGAGGCTCTTACAGTCTGATGCTTACCGTT 1456  
:|||||  
446 uTyrIleGlyValThrValGluGlyProPheLysSerAspHisTyrArgT 463  
1457 AT 1458  
||  
463 yr 463

seq\_name: SwissProt\_40:SAHH\_RHOSH

seq\_documentation\_block:  
ID SAHH\_RHOSH STANDARD; PRT; 463 AA.  
AC 050562;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine  
DE hydrolase) (AdoHcyase).  
GN AHCY.  
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
OC Rhodobacter.  
OX NCBI\_TaxID=1063;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97354111; PubMed=9210332;  
RA Mizoguchi H., Masuda T., Nishimura K., Shimada H., Ohta H., Shioi Y.,  
RA Takamiya K.  
RT "Nucleotide sequence and transcriptional analysis of the flanking  
RT region of the gene (spb) for the trans-acting factor that controls  
RT light-mediated expression of the puf operon in Rhodobacter  
RT sphaeroides.";  
RL Plant Cell Physiol. 38:558-567(1997).  
CC -|- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine  
CC + L-homocysteine.  
CC -|- COFACTOR: NAD (BY SIMILARITY).  
CC -|- PATHWAY: ACTIVATED METHYL CYCLE.  
CC -|- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; U76671; AAB88245.1; -  
DR HSP: P10760; 1B3R.  
DR InterPro: IPR000043; Ado\_hcyase.  
DR Pfam: PF00670; AdoHcyase; 1.  
DR PROSITE: PS00738; ADHCHYASE\_1; 1.  
DR PROSITE: PS00739; ADHCHYASE\_2; 1.  
KW Hydrolase; NAD: One-carbon metabolism.  
FT NP\_BIND 247 278 NAD (POTENTIAL).  
SEQUENCE 463 AA; 50617 MW; ACE8EB9D0D0CCBBA CRC64;

alignment\_scores:  
Quality: 1359.00 Length: 487  
Ratio: 3.634 Gaps: 8  
Percent Similarity: 76.797 Percent Identity: 58.932  
alignment\_block:  
US-09-759-990-1 x SAHH\_RHOSH ..  
Align seg 1/1 to: SAHH\_RHOSH from: 1 to: 463  
34 GAGTACAGAAATGCGACATCAACATCATCTGTTCTCGCGCGTAAGAACT 83  
:|||||  
3 AspPheIleValLysAspLeuSerLeuAlaAspPheGlyArgLysGluLe 19  
84 TACCCTTCTGAGAGAGAAATCCAGGCTTATGTTCTTCGTGACCGTT 133  
1:|||||  
19 uAspIleAlaGluThrGluMetProGlyLeuMetAlaLeuArgGluGluP 36  
134 ATTCGCTTCTTAAGCCATTGAAGGTTGTCAGATCTCTGTTCCCTCCAC 183  
:|||||  
36 heGlyAlaSerLysProLeuLysGlyAlaArgIleAlaGlySerLeuHis 52  
184 ATGACAGTCCAGACAGCCCTCTCATCGAGACACTCACAGCTCTTGTGC 233  
|||||  
53 MetThrValGlnThrAlaValLeuIleGluThrLeuValAlaLeuGlyAl 69  
234 TGATGTCAGATGGCTTCTTCAACATCTTCTACACAGATACACCGG 283  
|||||  
69 aspValArgTyrAlaSerCysAsnIlePheSerThrGlnAspHisAla 86  
284 CTGCTCTATCTGTTGTCGCCCAACAGGCACACAGAGAAAGCCAGCCGT 333  
:|||||  
86 laAlaAlaIleAlaAla.....SerGly 93  
334 ATCCCATGCTTCTGCGCTGGAAGGCGAAACACTCCAGAAATACAGGAGAA 383  
:|||||  
94 ValProValPheAlaIleLysGlyGluThrLeuGluAspTyrTrpAlaTy 110  
384 CACATACCGCGCTCTCATCGCCGATGTCAGGCCACAGCCACAGAGTTG 433  
110 rThrAspLysIlePheGlnPheAlaAspGlyThr...CysAsnMetIleL 126  
434 TCATGATGTTGGTGATGCTACACTCTCTCATCTCAAGGCTTCGAATTC 483  
126 euAspAspGlyGlyAspAlaThrLeuTyrIleLeuLeuGlyAlaArgVal 142  
484 GAACAGCCGCT.....GCTGTTCCAGAGCCCAACAGAGAAGC 518  
143 Glu...AlaGlyGluThrAspLeuIleAlaValProGlnSerAspGluG1 158  
519 TGCAACCTCGAATACCGCTCGGTCTTGTCTACACTCAAGCAGGCTTCA 568  
158 uValCysLeu.....PheA 163  
569 ACCAAGCAAGAACAC.....TGGCACACAGTTGCT 600  
163 snGlnIleArgLysArgMetAlaGluThrProGlyTyrPheThrLysGln 179  
601 GCC...GCATGAAGGTTGTTCCGAAGAGACAAACAGAGTGTCCACCG 647  
180 ArgAspAlaIleLysGlyValSerGluGluThrThrGlyValHisar 196  
648 CCTCTACACAGTTCGAGAGGAGGCAAACTCTCTTCCAGCCATCAACG 697  
196 gLeuTyrAspLeuHisLysLysGlyLeuLeuProPheProAlaIleAsnV 213  
698 TCAACGACGCTGTTACAAAGTCCAAAGTTCGATAACATCTACGGTCCCGC 747  
213 aLAsnAspSerValThrLysSerLysPheAspAsnLysTyrGlyCysLys 229  
748 CACTCCCTTATCGATGATCAACCGCTGTCCTCGGATGTGATGATCGCGG 797

```

.....
230 GuSerLeuValAspGlyIleArgAlaThrAspValMetMetAlaGl 246
798 CAAGACAGCTCTCGTATGATGACGGCGATGTCGCAAGGGCTCGCGTC 847
|||||.....
246 yLysValAlaValValCysGlyTyrGlyAspValGlyGlySerAlaA 263
848 ATCCCTCGTGGCCAGGCGTCGGTTCATATCATCATCAGAACTCGACCA 897
|||||.....
263 laSerLeuArgGlyAlaGlyAlaArgValLysValThrGluValAspPro 279
898 ATCTCGCTCTCCAGCTGCTCATGGAAGGCTACAGGTCGCGCCATCGA 947
|||||.....
280 ileCysAlaLeuGlnAlaMetaspGlyPheGluValValValLeuGl 296
948 GGAATCGTCAAGGATGTCATATCTTCGTATCATGCACAGGAACCTCGG 997
|||||.....
296 uAspValValGlnAspAlaAspIlePheIleThrThrGlyAsnArgA 313
998 ATATCATCTCTTTCATGATGATGCGCCAGATGAAGTAAGGCTATGTC 1047
|||||.....
313 spValIleArgIleGluHisMetArgGluMetLysAspMetAlaIleVal 329
1048 GGTACATCGGCACTTCGATACGAAATGATACAGATCGCTCATGAA 1097
|||||.....
330 GlyAsnIleGlyHisPheAspAsnGluIleGlnValAlaAlaLeuLysAs 346
1098 ATACCCAGGATCAAGCACATCCCAATCAAGCGCAATACGACATGTGGG 1147
|||||.....
346 nHis.....LysTrpThrAsnIleLysAspGlnValAspMetIleG 360
1148 AATCCAGATGGCCAGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1197
|||||.....
360 luMetProSerGlySerArgIleIleLeuLeuSerGluGlyArgLeuLeu 376
1198 AACCTTGCTCGCTTACAGGTACAGGTACAGGTACAGGTACAGGTACAG 1247
|||||.....
377 AsnLeuGlyAsnAlaThrGlyHisProSerPheValMetSerAlaSerPh 393
1248 CACAACACGACATCGCTACGCTACGCTACGCTACGCTACGCTACGCTAC 1297
|||||.....
393 eThrAsnGlnValLeuAlaGlnIleGluLeuTrpThrLysGlyAlaAspT 410
1298 TCGAGAAGAAGGTATACACATCTCCGACATCTCGATGATGAAGTCGCT 1347
|||||.....
410 yrGlnProGlyValTyrIleLeuProLysAlaLeuAspGluLysValAla 426
1348 CGCTCCACCTCGGATCTCTCGATCTCCACCTTACAAAGCTTACACAGAA 1397
|||||.....
427 ArgLeuHisLeuLysLysIleGlyValLysLeuThrAspValArgProGl 443
1398 GCAGCTGACTTACATCACTTCCAGTTGAGGTCCTTACAGTCTGATG 1447
|||||.....
443 uGlnAlaAspTyrIleGlyValLysValGluGlyProPheLysAlaGluH 460
1448 CTTACCGTTAT 1458
|||||.....
460 lstyArgTyr 463
seq_name: SwissProt_40:SAHH_PLAF7
seq_documentation_block:
ID SAHH_PLAF7 STANDARD; PRT; 479 AA.
AC P50250;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
OS Plasmidium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.
OX NCBI_TaxID=36329;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=94266832; PubMed=8206944;
RA Creeson K.A., Rathod P.K., Wellens T.E.;
RT "Plasmodium falciparum S-adenosylhomocysteine hydrolase. cDNA
RT identification, predicted protein sequence, and expression in
RT Escherichia coli.";
RL J. Biol. Chem. 269:16364-16370(1994).
CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -!- COFACTOR: NAD (BY SIMILARITY).
CC -!- PATHWAY: ACTIVATED METHYL CYCLE.
CC -!- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U07365; AAA21391.1; --
DR HSP; P10760; IB3R.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF00670; AdoHcyase; 1.
DR PROSITE; PS00738; ADHCHYASE_1; 1.
DR PROSITE; PS00739; ADHCHYASE_2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
FT NP_BIND 258 289 NAD (POTENTIAL).
SQ SEQUENCE 479 AA; 53893 MW; 74F4D754E6F8970A CRC64;

```

## alignment\_scores:

```

Quality: 1332.00 Length: 481
Ratio: 3.469 Gaps: 5
Percent Similarity: 79.834 Percent Identity: 53.430

```

## alignment\_block:

```

US-09-759-990-1 x SAHH_PLAF7

```

```

Align seg 1/1 to: SAHH_PLAF7 from: 1 to: 479

```

```

40 AGAATGCCGACATCAACCTCCATGTTCTCGCGCGTAGGAACCTTACCCT 89
:|||||.....
7 LysValLysAspIleSerLeuAlaProPheGlyLysMetGlnMetGluIl 23
90 TGTCTGAGAAAGAAATGCCAGGTCTTATGTTCTTCGTGAGCGTTATTCGG 139
:|||||.....
23 eSerGluAsnGluMetProGlyLeuMetArgIleArgGluGluGlyGlyL 40
140 CTTCTAAGCCATTGAAGGGTGTGAGAACTCTGTTCCCTCCACATGACA 189
:|||||.....
40 ysAspGlnProLeuLysAsnAlaLysIleThrGlyCysLeuHisMetThr 56
190 GTCCAGACAGCGCTCCTCATCGAGACACTCACAGCTCTTGGTGTGATGT 239
|||||.....
57 ValGluCysAlaLeuLeuIleGluThrLeuGlnLysLeuGlyAlaGlnIl 73
240 CAGATGGGCTTCTCGCAACATCTTCTTACACAAAGATACAGCGCTGCTG 289
:|||||.....
73 eArgTrpCysSerCysAsnIleTyrSerThrAlaAspTyrAlaAlaAla 90
290 CTATCGTGTCTCGCCCAACAGGCACACAGAGAGCCAGCCGGTATCCCA 339
|||||.....
90 laVal.....SerThrLeuGlu.....AsnValThr 98
340 GTCTTCGCTGGAAGGGCAACACTCTCCCAAGATACTGGGAGAACACATA 389
|||||.....

```

99 ValPheAlaTrpLysAsnGluThrLeuGluGluTyrTrpTrpCysValG1 115  
390 CCGCGCTCTACATGGCCAGATGTT ..... CAAGGCCACACAGAGGTG 433  
115 uSerAlaLeuThrTrpGlyAspGlyAspAsnGlyProAspMetIleV 132  
434 TCAGATGATGGTGTGATGCTACATCTCTCATCTCCAAAGGGCTTCGAA... 480  
132 AlaAspAspGlyGlyAspAlaThrLeuLeuValHisGlyValGluTyr 148  
481 .....TTCCGAACACAGCCGGTCTGCTTCAGAGCCACACAGAGCTGA 521  
149 GluLysLeuTyrGluGluLysAsnIleLeuProAspProGluLysAlaL 165  
522 CAACCTCGAATACCGCTCGTCTTCTGCTACACTCAAGCAGAGGCTTCAAC 571  
165 sAsnGluGluGluArgCysPheLeuThrLeuLeuLysAsnSerIleLeuL 182  
572 AAGACAAGAACCACTGGCACACAGTGTGCTGCGGATGACAGCGTGTTC 621  
182 yAsnProLysLysTrpThrAsnIleAlaLysLysIleIleGlyValSer 198  
622 GAAGAGACAACAGAGGTGTCACCGCTCTACAGCTCGAGAGGAGG 671  
199 GluGluThrThrThrGlyValLeuArgLeuLysLysMetAspLysGln 215  
672 CAACCTCTCTTCCAGCCATCAAGCTCAAGCAGCTGTACAAAGTCCA 721  
215 nGluLeuLeuPheThrAlaIleAsnValAsnAspAlaValThrLysGlnL 232  
722 AGTTCGATACATCTACGCTGCTGCGCCACTCTCTATCATCGATGATCAAC 771  
232 yTyrAspAsnValTyrGlyCysArgHisSerTyrThrAspGlyLeuMet 248  
772 CGTCTCCGATGTCATGATCGCGCGCAGACAGCTCTGTCATGGTTA 821  
249 ArgAlaThrAspPheLeuIleSerGlyLysIleValIleCysGlyT 265  
822 CGGCGATCTCGCAAGGGTGGCTCAATCCCTCCGTCGCGCAAGGGCTC 871  
265 rGlyAspValGlyLysGlyCysAlaSerSerMetLysGlyLeuGlyAla 282  
872 CGGTATCATCAGAACTCGACCAATCTGGCTCTCAGGCTGCCATG 921  
282 rGlyValThrThrGluIleAspProIleCysAlaIleGlnAlaValMet 298  
922 GAAGCTACCAAGTCCGCGCATCGAGGAAGTCTCAAGGATCTCGATAT 971  
299 GluGlyPheAsnValThrLeuAspGluIleValAspLysGlyAspPh 315  
972 CTTCGTTATCATCAGAGAACTCGCATATCATCTCTGTTGACATGATGG 1021  
315 ePheIleThrCysThrGlyAsnValAspValIleLysLeuGluHisLeuL 332  
1022 CCCAGATGAGGATAGGATATGTCGTTAATCATCGGCCACTTCGATAAC 1071  
332 eUlysMetLysAsnAlaValAlaValGlyAsnIleGlyHisPheAsp 348  
1072 GAAATTTGATACAGTGGCTCTCATGAATACCCAGGCATCAAGCACATCCC 1121  
349 GluIleGlnValAsnGluLeuPheAsnTyrLysGlyIleHisIleGluAs 365  
1122 AATCAAGCCAGAAATACGACATGTTGGAAATCCAGATGGCCACGCTATCC 1171  
365 nValLysProGlnValAspArgIleThrLeuProAsnGlyAsnLysIleI 382  
1172 TCCTTCTCTGAGGCGCCCTCTTAACTTCACTGCTGCTGCTGCTGCTGCT 1221  
382 lValLeuAlaArgGlyArgLeuLeuAsnLeuGlyCysAlaThrGlyHis 398  
1222 CCATCTTCTGTTATGTCATGTCATTCACAAACACAGACACTCGCTCAGCT 1271  
399 ProAlaPheValMetSerPheCysAsnGlnThrPheAlaGlnLe 415

1272 CGACCTCTACGAAAGAGA.....GGAAATCTCGAGAAAGGTTTACA 1315  
415 uAspLeuTrpGlnAsnLysAspThrAsnLysTyrGluAsnLysValTyrL 432  
1316 CACTTCCGAAGCATCTCGATGAGAGAGTGCCTCCACCTCGGATCT 1365  
432 euLeuProLysHisLeuAspGluLysValAlaLeuTyrHisLeuLysLys 448  
1366 CTCGATGTCCACCTTACAAAGCTTACACAGACAGAGCTGACTACATCAA 1415  
449 LeuAsnAlaSerLeuThrGluLeuAspAsnGlnCysGlnPheLeuG1 465  
1416 GCTTCCAGTTGAGGTCCTTACAAAGTCTGATGCTTACCGTTAT 1458  
465 yValAsnLysSerGlyProPheLysSerAsnGluTyrArgTyr 479

seq\_name: SwissProt\_40:SAHH\_RAT

seq\_documentation\_block:

ID SAHH\_RAT STANDARD; PRT; 431 AA.  
AC P10760;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine  
hydrolase) (AdoHcyase).  
GN AHCY.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=87118240; PubMed=3027698;  
RA Ogawa H., Gomi T., Mueckler M.M., Fujioaka M., Backlund P.S. Jr.,  
RA Aksamit R.R., Unson C.G., Cantoni G.L.;  
RT "Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from rat  
liver as derived from the cDNA sequence.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:719-723(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FTSCHER 344;  
RX MEDLINE=95262723; PubMed=7744082;  
RA Merta A., Aksamit R.N., Kasir J., Cantoni G.L.;  
RT "The gene and pseudogenes of rat S-adenosyl-L-homocysteine  
hydrolase.";  
RL Eur. J. Biochem. 229:575-582(1995).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RC TISSUE=Liver;  
RX MEDLINE=99315209; PubMed=10387078;  
RA Hu Y., Komoto J., Huang Y., Gomi T., Ogawa H., Takata Y., Fujioaka M.,  
RA Takusagawa F.;  
RT "Crystal structure of S-adenosylhomocysteine hydrolase from rat  
liver.";  
RL Biochemistry 38:8323-8333(1999).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY. (2.8 ANGSTROMS).  
RX MEDLINE=20493601; PubMed=10913437;  
RA Komoto J., Huang Y., Gomi T., Ogawa H., Takata Y., Fujioaka M.,  
RA Takusagawa F.;  
RT "Effects of site-directed mutagenesis on structure and function of  
recombinant rat liver S-adenosylhomocysteine hydrolase. Crystal  
structure of D244E mutant enzyme.";  
RL J. Biol. Chem. 275:32147-32156(2000).  
CC -1- FUNCTION: ADENOSYLOHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF  
S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;  
THEREFORE ADENOSYLOHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE  
CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR  
CONCENTRATION OF ADENOSYLOHOMOCYSTEINE.  
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine  
+ L-homocysteine.

CC -I- COFACTOR: NAD.  
 CC -I- PATHWAY: ACTIVATED METHYL CYCLE.  
 CC -I- SUBUNIT: HOMOTETRAMER.  
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -I- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M15185; AAA40705.1; -;  
 DR EMBL; U14937; AAA92043.1; -;  
 DR PIR; A26583; A26583.  
 DR PDB; 1B3R; 23-DEC-98.  
 DR PDB; 1D4F; 17-JAN-01.  
 DR PDB; 1D4G; 17-JAN-01.  
 DR InterPro; IPR000043; Ado\_hcyase.  
 DR Pfam; PF00670; AdoHcyase; 1.  
 DR PROSITE; PS00738; ADHOCYASE\_1; 1.  
 DR PROSITE; PS00739; ADHOCYASE\_2; 1.  
 KW Hydrolase; NAD; One-carbon metabolism; 3D-structure.  
 FT INIT\_MET 0  
 FT NP\_BIND 213 244 NAD (POTENTIAL).  
 SQ SEQUENCE 431 AA; 47407 MW; 537154C9EAA4E380 CRC64;

## alignment\_scores:

Quality: 1286.00 Length: 478  
 Ratio: 3.695 Gaps: 5  
 Percent Similarity: 72.803 Percent Identity: 56.695

## alignment\_block:

US-09-759-990-1 x SAHH\_RAT ..

Align seg 1/1 to: SAHH\_RAT from: 1 to: 431

37 TACAGAAATGCGGACATCAACCTCCATGTTCTCGGCGTAAGAACTTAC 86  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 6 TyrLysValAlaAspIleGlyLeuAlaIaIaTrpGlyArgLysAlaLeu 22  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 87 CCTTCTCAGAGGAAATGCCAGGTCTTATGTTCTTCGTAGCGTTATT 136  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 22 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgGluMetTyr 39  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 137 CCGTTTCTAAGCCATTGAAGGTGTGAGAACTCTCTGTTCCCTCCACATG 186  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 39 erAlaSerLysProLeuLysGlyAlaIaIaGileAlaGlyCysLeuHisMet 55  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 187 ACAGTCAGACAGCGCTCTCATCGAGACACTCACAGCTCTTGGTCTGA 236  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 56 ThrValGluThrAlaValLeuIleGluThrLeuValAlaLeuGlyAlaG 72  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 237 TGTTCAGATGGCTTCTTCCACATCTTCTTACACAGATACAGCGGTG 286  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 72 uValArgTrpSerCysAsnIlePheSerThrGlnAspHisAlaAla 89  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 287 CTGCTATCGTTGTGGGCCCAACAGGCACACAGAGAGCCAGCGGTATC 336  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 89 laAlaIleAla.....lysAlaGlyIle 96  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 337 CCAGTCTTCGCTGGAGGGGAAACACTCCCAATATCTGGGAAACAC 386  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 97 ProValPheAlaTrpLysGlyGluThrAspGluGluTyrLeuTyrCys 113  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 387 ATACCGCGCTCTCACATGTCGAGATGTTCAGGCCCCACAGAGGTTCG 436  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 113 eGluGlnThrLeuHisPheLysAspGlyPro...LeuAsnMetIleLeu 129  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 437 ATGATGGTGTGATGTACACTCCTCATCTCTCAAGGCGTTCGAATTCGA 486

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 129 spAspGlyGlyAspLeuThrAsnLeuIle..... 138  
 487 ACAGCGGTGCTGTCTCCAGAGCCAAACAGAGCTCACAACTCGAATACCG 536  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 138 ..... 138  
 537 CTGCGTTCTTGTACTACTCAAGCAGGTCTTCAACCAAGACAAGAACCACT 586  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 138 ..... 138  
 587 GGCACACA.....GTTGCTGCCGGCATGACGGTGTTCCTCGAA 624  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 139 ..HisThrLysHisProGlnLeuLeuSerGlyIleArgGlyIleSerGlu 154  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 625 GAGACAACAACAGGTGTCCACCGCTCTTACAGCTCGAGAAGAGGGCAA 674  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 155 GluThrThrThrGlyValHisAsnLeuTyrLysMetMetAlaAsnGlyI 171  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 675 ACTCCTCTTCCAGCCATCAAGCTCAACGACGCTGTTCACAAAGTCCAAGT 724  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 171 eLeuLysValProAlaIleAsnValAsnAspSerValThrLysSerLys 188  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 725 TCGATAACATCTACGGCTCGCGCCCACTCCCTTATCGATCGTATCAACCGT 774  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 188 heAspAsnLeuTyrGlyCysArgGluSerLeuIleAspGlyIleLysArg 204  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 775 GCTTCCGATGTCATGTCGCGGCGGCAAGACAGCTCTCGTATCGGTTACGG 824  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 205 AlaThrAspValMetIleAlaGlyLysValAlaValAlaGlyTyrG 221  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 825 CGATGTCGCAAGGCTCGCTCAATCCCTCGTGGCCCAAGCGCTCGCG 874  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 221 YAspValGlyLysGlyCysAlaGlnAlaLeuArgGlyPheGlyAlaArg 238  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 875 TTATCATCACAGAACTCGACCAATCTCGCTCTCCAGGCTGCCATGAA 924  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 238 alIleIleThrGluIleAspProIleAsnAlaLeuGlnAlaAlaMetGlu 254  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 925 GGCTACAGGTCCGCCCATCGAGAAAGTCGTCAGAGATGTCGATATCT 974  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 255 GlyTyrGluValThrThrMetAspGluAlaCysLysGluGlyAsnIle 271  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 975 CGTTACATGACAGGAAACTCGCATATCATCTCTGTTGACATGATGGCC 1024  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 271 eValThrThrThrGlyCysValAspIleIleLeuGlyArgHisPheGlu 288  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 1025 AGATGAAGATAGGCTATGTCGGTAACATCGGCCACTTCGATAACGAA 1074  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 288 lMetLysAspAlaIleValCysAsnIleGlyHisPheAspValGlu 304  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 1075 ATTGATACAGATGGCTCATGAATACCCAGGATCAAGCACATCCCCAAT 1124  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 305 IleAspValLysTrpLeuAsnGlu...AsnAlaValGluLysValAsn 320  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 1125 CAAGCCAGATACGATGTCGGAATCCAGATGGCCAGCTATCTCTCC 1174  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 320 eLysProGlnValAspArgTyrLeuLeuLysAsnGlyHisArgIleIle 337  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 1175 TTCTTGTCTGAGGCGCTCTTAACCTTGGCTGCGCTACAGGTACCCCA 1224  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 337 euLeuAlaGluGlyArgLeuValAsnLeuGlyCysAlaMetGlyHisPro 353  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 1225 TCTTTCGTTATGTCATGTCATCAAAACACAGACACTCGCTCAGCTCGA 1274  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 354 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIle 370  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 1275 CCTCTACGAAAAAGAGAGAAATCTCGAGAGAGGTTTACACTTCCGA 1324  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 370 uLeuTrpThrHisProAspLysTyrProValGlyValHisPheLeuPro 387  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 1325 AGCATCTCGATGAAGACTCGCTCGCTCCACTCGGATCTCTCGATGTC 1374  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

us-09-759-990-1-n2p.rsp

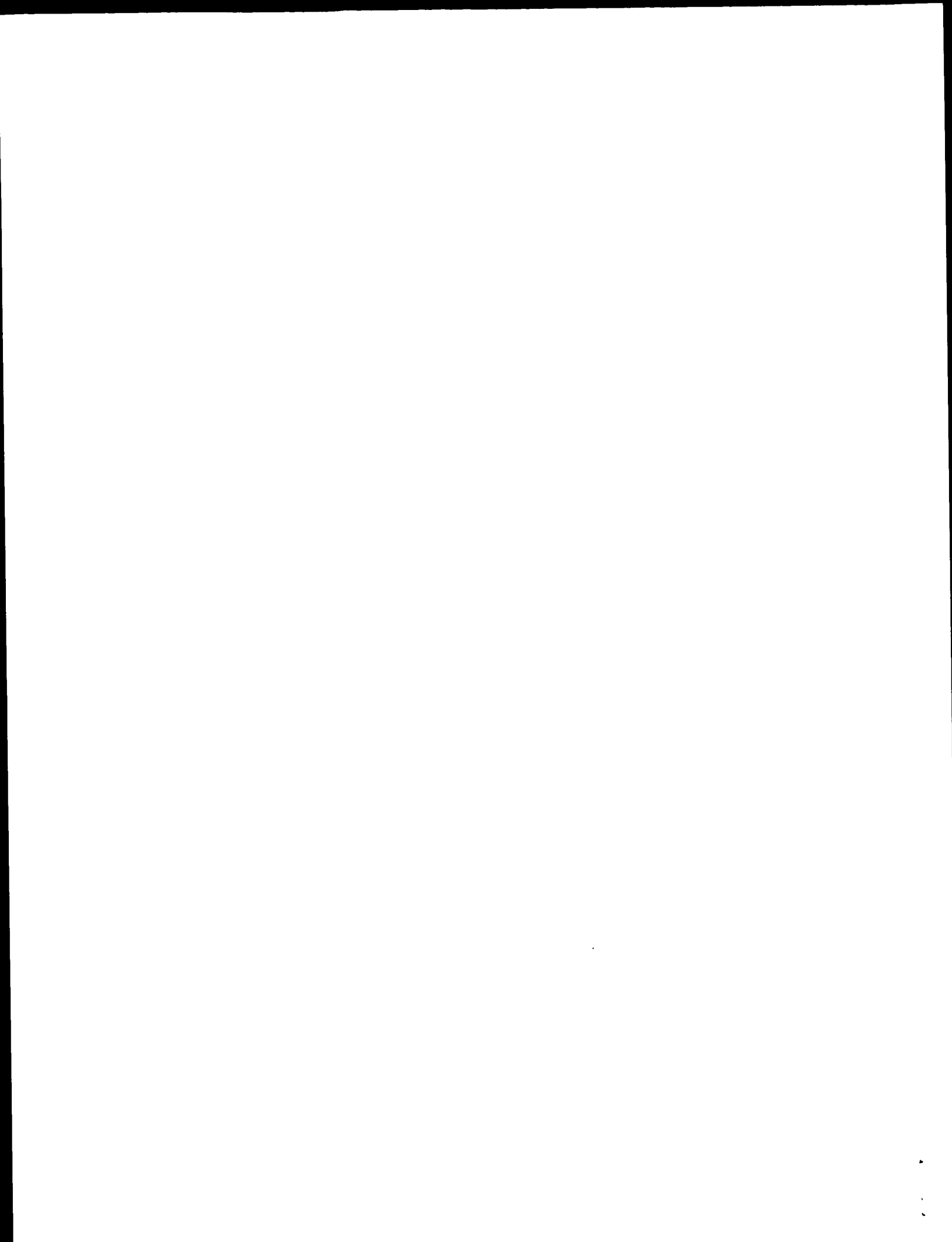
Mon Aug 19 15:16:21 2002

```

387 ysLysLeuAspGluAlaValAlaGluAlaHisLeuGlyLysLeuAsnVal 403
1375 CACCTTCAAGCTTACACAGCAGCGTCTGACTACATCAACGTTCCACT 1424
|||||
404 LysLeuThrLysThrGluGlyGlnAlaGlnTyrLeuGlyMetProI 420
1425 TGAGGGTCTTACAGTCTGATGCTTACCGTTAT 1458
420 eAsnGlyProHeLysProAspHisTyrArgTyr 431
seq_name: SwissProt_40:SAHH_MOUSE
seq_documentation_block:
ID SAHH_MOUSE STANDARD; PRT; 431 AA.
AC P50247; 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase) (Liver copper binding protein) (CUBP).
GN AHCY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
STRAIN-BALB/C; TISSUE=Liver;
MEDLINE=95386522; PubMed=7657650;
RA Bethin K.E., Petrovic N., Rittinger M.J.;
RT Identification of a major hepatic copper binding protein as S-
RT adenosylhomocysteine hydrolase.
RL J. Biol. Chem. 270:20698-20702(1995).
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
CC -| CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -| COFACTOR: NAD.
CC -| PATHWAY: ACTIVATED METHYL CYCLE.
CC -| SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -| SUBCELLULAR LOCATION: Cytoplasmic.
CC -| SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC
CC EMBL: L32836; AAA70378.1; -.
CC DR HSPSP; P10760; I1B3R.
CC DR MGD; MGI:87968; Ahcy.
CC DR InterPro: IPR000043; Ado_hcyase.
CC DR Pfam: PF00670; AdoHcyase; 1.
CC DR PROSITE; PS00738; ADHCYASE_1; 1.
CC DR PROSITE; PS00739; ADHCYASE_2; 1.
CC KW Hydrolase; NAD; One-carbon metabolism.
CC FT INIT_MET 0
CC FT NP_BIND 213 244 NAD (POTENTIAL).
CC SEQUENCE 431 AA; 47543 MW; D57DCE9D98C9B5 CRC64;
alignment_scores:
Quality: 1278.00 Length: 478
Ratio: 3.662 Gaps: 5
Percent Similarity: 73.013 Percent Identity: 56.485
alignment_block:
US-09-759-990-1 x SAHH_MOUSE
Align seg 1/1 to: SAHH_MOUSE from: 1 to: 431
37 TACAGATATCCGACATCAACCTCCATGTTCTCGCGGTGAAGACATTAC 86
|||||
6 TyrLysValAlaAspIleGlyLeuAlaIleAlaIleAlaIleAlaIle 22
87 CCTGCTGAGAGAAATGCCAGCTCTATGTTCTTCTGTCGAGCGTTATT 136
|||||
22 PileAlaGluAsnGluMetProGlyLeuMetArgMetArgGluMetTyrS 39
137 CCGCTTCTAAGCCATTGAAGGGTGTCAAGATCTCTGTTTCCCTCCACATG 186
|||||
39 eAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 55
187 ACAGTCAGACAGCGCTCTCATCGAGACACTCACAGCTCTTGGTGTGA 236
|||||
56 ThrValGluThrAlaValLeuIleGluThrLeuValAlaLeuGlyAlaG 72
237 TGTCAAGATGGCTTCTGCAACATCTTCTTACACAAGATACAGCGCTG 286
|||||
72 uValArgTrpSerCysAsnIlePheSerThrGlnAspHisAlaAla 89
287 CTGCTATCGTTGTCGCCCAACAGCACACAGAGCCAGCGGTATC 336
|||||
89 laAlaIleAla.....LysAlaGlyIle 96
337 CCAGTCTTCGCTGGAAGGGCAACACTCCAGATATCTGGGAGAACAC 386
|||||
97 ProValPheAlaTrpLysGlyGluThrAspGluGluTyrLeuTrpCys 113
387 ATACCGGCTCTCATGATGCCAGATGGTCAAGCCACACAGCGGTGTGCG 436
|||||
113 eGluGlnThrLeuHisPheLysAspGlyPro...LeuAsnMetIleLeu 129
437 ATGATGGTGGTGTGATGTCATCTCTATCTCCAAAGGCTTCGAATTGAA 486
|||||
129 sPaspGlyLysAspLeuThrAsnLeuIle..... 138
487 ACAGCGGGTGTGTTCCAGAGCCACAGAGAGTGCACACCTCGNATACCG 536
138 ..... 138
537 CTGCGTTCCTGCTACACTCAAGCAGGCTTCAACAGACAGAACCACT 586
138 ..... 138
587 GGCACACA.....GTTGCTCCCGCATGACGCTGTTCCGAA 624
|||||
139 ..HisThrLysTyrProGlnLeuLeuSerGlyIleArgGlyIleSerGlu 154
625 GAGACACACACAGTGTCCACCGCTCTACAGCTCGAGAGAGGAGGCAA 674
|||||
155 GluThrThrThrGlyValHisAsnLeuTyrLysMetSerAsnGlyIle 171
675 ACTCCTCTTCCCGCATCAACGTCACAGCGTGTTCACAAAGTCCAAGT 724
|||||
171 eLeuAsnValProAlaIleAsnValAsnAspSerValThrLysSerLysP 188
725 TCGATACATCTACGCTGCGCGCCACTCCCTTATCGATGTTATCAACCGT 774
|||||
188 heAspAsnLeuTyrGlyCysArgGluSerLeuIleAspGlyIleLysArg 204
775 GCTTCCGATGTCATGATCGCGGCGCAAGACAGCTCTCTGTCGTTACGG 824
|||||
205 AlaThrAspValMetIleAlaGlyLysValAlaValAlaGlyTyrG 221
825 CGATGTCGGAAGGCTCGCTCAATCTCCGTCGTCGCAAGCGCTCGCG 874
|||||
221 yAspValGlyLysGlyCysAlaGlnAlaLeuArgGlyPheGlyAlaArg 238
875 TTATCATCACAGAACTCGACCAATCTCGCTCTCCAGGCTGCCATGGAA 924

```

```
|||||
238 aillelethrGluilleaspProilleasnAlalaLeuGlnAlalaMetGlu 254
925 GCCTACCGAGTCGCGCGCATCGAGGAAGTCGTCAAGGATGTCGATATCTT 974
|||||
255 GlyTyrGluValThrThrMetaspGluAlaCysLysGluGlyAsnIlePh 271
975 CGTTACATGCACAGGAACCTGCATATCATCTCTGTTGCATGATGGCCC 1024
|||||
271 evalthrThrGlycysValaspIleIleLeuGlyArgHisPheGluG 288
1025 AGATGAAGGATAAGGCTATTGTGCGTAACATCGGCCACTTCGATAACGAA 1074
|||||
288 lnmetyaspAlaIleValCysAsnIleGlyHisPheaspValGlu 304
1075 ATTGATACAGATGGCCTCATGAAATACCGAGGCATCAAGCACATCCCAAT 1124
|||||
305 IleaspValLysTrpLeuAsnGlu...AsnAlaValGluLysValAsnI 320
1125 CAAGCCAGAATACGACATGTGGGAATTCACAGATGCCACGCTATCCTCC 1174
|||||
320 eLysProGlnValaspArgTyrTrpLeuLysasnGlyArgGilelleL 337
1175 TTCTTGCTGAGGGCCGCTTCTTAACCTTGGCTGCGCTACAGGTCACCCA 1224
|||||
337 euLeuAlaGluGlyArgLeuValasnLeuGlyCysAlaMetGlyHisPro 353
1225 TCTTTCGTTATGTCATGTCATTCAACACAGACACACGCTCAGCTCGA 1274
|||||
354 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIleGl 370
1275 CCTCTACGAAAGAGAGGAAATCTCGAGAAGAAGGTTTACACACTTCCGA 1324
|||||
370 uLeuTrpThrHisProAspLysTyrProValGlyValHisPheLeuProL 387
1325 AGCATCTCGATGAGAAGTCGCTCGCTCCACCTCGGATCTCTCGATGTC 1374
|||
387 ysLysLeuaspGluAlaValAlaGluAlaHisLeuGlyLysLeuAsnVal 403
1375 CACCTTACAAAGCTTACACAGAAGCAGGCTGACTACATCAACGTTCCAGT 1424
|||||
404 LysLeuThrLysLeuThrGluLysGlnAlaGlnTyrLeuGlyMetProIl 420
1425 TGAGGGTCCTTACAAGTCGTGCTTACCGTTAT 1458
|||||
420 eAsnGlyProPheLysProAspHisTyrArgTyr 431
```







```

1328 ATCTCGATGAAGAGTCGCTCGCTCCACCTCCGATCTCTCGATGTCAC 1377
|||||
452 isLeuaspGluLysValAlaArgIleHisValGluAlaLeuGlyGlyHis 468
|||||
1378 CTTTACAAGGTTTACACAGACGAGCTGACTACATCAACGTTCCAGTTGA 1427
|||||
469 LeuThrLysLeuThrLysGluGlnAlaGluTyrLeuGlyValAspValGI 485
|||||
1428 GGGTCCTTACAAGTCTGATGCTTACCGTTAT 1458
|||||
485 uGlyProTyrLysProAspHisTyrArgTyr 495
|||||
seq_name: pir2:D87005

seq documentation_block:
probable S-adenosyl-L-homocysteine hydrolase [imported] - Mycobac
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 1
C;Accession: D87005
R;Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson,
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skel
A;Title: Massive gene deletion in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: D87005
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <SNO>
A;Cross-references: GB:AL450380; NID:gl3092881; PIDN:CAC30280.1;
C;Genetics:
A;Gene: sahH
C;Superfamily: adenosylhomocysteinase

alignment_scores:
Quality: 1537.50 Length: 477
Ratio: 3.932 Gaps: 3
Percent Similarity: 81.971 Percent Identity: 62.264

alignment_block:
US-09-759-990-1 x D87005 ..

Align seg 1/1 to: D87005 from: 1 to: 492

34 GAGTACAGAAATGCCGACATCAACCTCCATGTTCTCGCGCTAAGGAACT 83
|||||
17 AspPheLysValAlaAspLeuSerLeuAlaAsnPheGlyArgLysGluIle 33
|||||
84 TACCCTTCGTGAGAAGAAATGCCAGGTCTTATGGTCTTCGTGAGCGTT 133
|||||
33 uAspLeuAlaGluTyrGluMetProGlyLeuMetSerLeuArgHisGluT 50
|||||
134 ATTCCGCTTCTAAGCCATTGAAGGGGTGCAGAACTCTCTGTTCCCTCCAC 183
|||||
50 YrAlaGluValGlnProLeuLysGlyAlaArgIleSerGlySerLeuHis 66
|||||
184 ATGACAGTCCAGACAGCGCTCCTCATCCAGACACTCACAGCTCTTTGGTGC 233
|||||
67 MetThrValGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 83
|||||
234 TGATGTCAGATGGGCTTCCTCAACATCTTCTTACACAAGATACAGCGCG 283
|||||
83 aGluValArgTrpAlaSerCysAsnIlePheSerThrGlnAspHisAla 100
|||||
284 CTGCTGCTATCGTTGTCGGCCCAACAGGCACACACAGAGAACCCAGCGGT 333
|||||
100 laAlaAlaValValGlyProTyrGlyThrProGluGluProLysGly 116
|||||
334 ATCCAGTCTTCGCTCGAAGGGCGAAACACATCCCGAGATACTGGGAGAA 383
|||||

```

117 ValProValPheAlaTrpLysGlyGluThrLeuGluGluThrTrpAl 133  
384 CACATACCGCGCTCTCACATGGCCAGATGTCAGGCCCCA...CAGCAGG 430  
133 aAlaGluGlnMetLeuThrTrpProAspLysProValAsnMetI 150  
431 TTGTGCATGATGGTGGTGATGTACACTCTCTCATCTCCAGGGCTTCGAA 480  
150 leLeuAspAspGlyGlyAspAlaThrMetLeuValLeuArgGlyValGln 166  
481 TTCGAACACGCGCGTCTCTTCACAGAGCCCAACAAGACTGACACACCTCGA 530  
167 TyrGluLysAlaGlyValValProAlaGluValAspSerAlaGln 183  
531 ATACCGCTGCGTCTTGTCTACACTCAAGCAGGTCTTCAACCAAGACAAGA 580  
183 utrLysValPheLeuAsnLeuLeuArgLysArgPheGluThrAspLysG 200  
581 ACCACTGGCACACACTTGTCTGCGCGCATCAACGCTGTTTCCGAAGAGACA 630  
200 lYlYsTrpThrLysIleAlaLysSerValLysGlyValThrGluGluThr 216  
631 ACACAGGTGTCCACCGCTCTACACAGCTCGAGAGAGGGCAACACTCCT 680  
217 ThrThrGlyValLeuArgLeuTyrGlnPheAlaAlaAalagLysAspLeuAl 233  
681 CTTCCACGCCATCAACGTCAACGACGCTCTTACAAAGTCCAAGTTCGATA 730  
233 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 250  
731 ACATCTAGCGGTGCGCGCACCTCCCTTATCGATGGTATCAACCGTGTCTCC 780  
250 snLysTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 266  
781 GATCTCATGATCGCGCGCAACACACTCTCTGTCATGGTTCACGGCATGT 830  
267 AspSerLeuIleGlyGlyAsnValLeuIleCysGlyTyrGlyAspVa 283  
831 CGGCAAGGCTCGCCTCAATCCCTCCCTGGCCAAAGCGCTCGCGTGTATCA 880  
283 lGlyLysGlyCysAlaGluAlaAlaLysGlyGlnGlyAlaArgValThrI 300  
881 TCACAGAACTCGACCCCAATCGCCTCTCCAGGCTGCCATGGAGGCTAC 930  
300 leThrGluIleAspProIleAsnAlaLeuGlnAlaLeuMetGluGlyPhe 316  
931 CAGGTGCGCGCATCGAGGAAGTGGTCAAGGATGTCGATATCTTCGTTAC 980  
317 AspValLysArgValGluAspValIleAlaAspSerAspIleValValTh 333  
981 ATCCACAGGAACCTGCCATCATCTCTGTTCACATGATGGCCAGATGA 1030  
333 rAlaThrGlyAsnLysAspIleIleLeuLeuGluHisMetLysAlaMetI 350  
1031 AGGATAGGCTATTGTTCGGTAACATCGGCCACTTCGATAACCAAAATTGAT 1080  
350 ysAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 366  
1081 ACAGATGGCCTCATGAATACCCAGGCATCAAGCACATCCCAATCAAGCC 1130  
367 MetAlaAlaLeuGluArg...SerGlyAlaThrArgLeuAsnIleLysPr 382  
1131 AGAATACGACATGTGGAAATCCAGAT...GGCCACGCTATCCTCCCTC 1177  
382 oGlnValAspLeuTrpThrPheGlyAspSerGlyLysSerIleIleValL 399  
1178 TTGCTGAGGGCGCCCTTCTTAACCTTCGGCTCGCGTACAGGTCAACCATCT 1227  
399 euSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 415  
1228 TTCGTTATCTCAATGTCAATTCACAAACCACACTCGCTAGCTCGACCT 1277  
416 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 432

[illegible]

8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24

```

69 CGCCGTAAGAACTTACCTTGTGAGAGGAATGCCAGGCTCTATGG 118
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMetA 41
119 TTTCTTGTGAGCGTTATTCGCTTCTAAGCCATTGAAGGGTGTCAAGATC 168
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57
169 TCTGTTCCTCCACATGACAGTCCAGACACCGCTCCTCATCGACACT 218
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
219 CACAGCTCTGGTCTGATGTCAGATGGCTTCTCGCAACATCTTCTCTA 268
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91
269 CACAAGATACAGCCGCTGCTCTGCTATGCTGTGCGCCCAACAGGCACCA 318
91 hrGlnAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319 GAGAAGCCAGCGGTATCCCAAGTCTTCCGCTGGAAGGGCGAACAACACTCC 368
106 .....ValPheAlaTrpLysGlyGluThrLeuG 115
369 AGAATCTGGGAAACACATACCGCGCTCTACATGGCCAGATGGTCAAG 418
115 nGluTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 132
419 GCCACAGCAGGTGTTTCGATGATGTTGATGCTGATGCTACATCTCTCATCTCC 468
132 lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148
469 AAGGCTTCGAA.....TTCGAAACAGCGCGTCTTCCAGA 506
149 GluGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProAs 165
507 GCCAACAGAGTCTGACAACTCGAATACCGCTGCTGCTGCTGCTGCTACACTCA 556
165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIleL 182
557 AGCAGGTCTTCAACCAAGACAGACACCATGGCCACACATGGTGTGCGCGC 606
182 ysGluGlyLeuGlnValAspProLysLysThrHisLysMetLysGluArg 198
607 ATGAACGCTGTTTCCGAGACACAAACAGGTGTCCACCGCCCTCTACCA 656
199 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyrgL 215
657 GCTCAGAGAGGAGGCAAACTCTCTTCCAGCCATCAAGCTCAACGACG 706
215 nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 232
707 CTGTACCAAGTCCAGTTCGATGATGATGATGATGATGATGATGATGATGAT 756
232 erValThrLysSerLysPheAspAsnLeuTyrgLysArgHisSerLeu 248
757 ATCGATGATGATCAACCGCTGCTTCCGATGATGATGATGATGATGATGATGAT 806
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
807 TCTGTCATGGTTCAGGCGATGTCGCGAGGCTGCGCTCAATCCCTCC 856
265 aValIleCysGlyTyrgLysAspValGlyLysGlyCysAlaAlaMetL 282
857 GTGGCCAAAGCGCTCGCTGATATCATCAAGTTCGACCAATCTGGCT 906
282 ysThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
907 CTCACGCTGCCATGGAAGCTACAGGTCGCGCATCGAGGAGTCTGT 956
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 315
```

```

957 CAAGGATGTCGATATCTTCGTTACATGCACAGGAACACTCGGATATCATCT 1006
315 lSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
1007 CTGTTGACATGATGCCCCAGATGAAGGATAAGGCTATTGTCGGTAACATC 1056
332 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
1057 GCCCACTTCGATAACGAAATTCATGATGCGCTCATGAAATACCCAGG 1106
349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTrpProG 365
1107 CATCAAGCACATCCCAATCAACAGCCAGATACGACATGTGGGAATCCAG 1156
365 yValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProG 382
1157 ATGGCCAGCT...ATCCCTCTTCTTCTGAGGCGCGCTTCTTAACTT 1203
382 luThrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
1204 GGCTCGCTACAGGTACCCATCTTTCGTTATGTCATGTCATTCACAAA 1253
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAATC 1297
415 nGlnValIleAlaGlnLeuGluLeuTrpAsnGluLysAlaSerGlyLysT 432
1298 TCGAAGAAGGTTTACACACTTCCGAAGCATCTCGATGAAGAGTCTGCT 1347
432 yGluLysLysValTyrgLysValLeuProLysHisLeuAspGluLysValAla 448
1348 CCGCTCCACTCGGATCTCGATGCTCCACTTACAAAGCTTACACAGAA 1397
449 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAs 465
1398 GCAGGCTGACTATCAACGTTCCAGTTGAGGTCCTTACAAAGTCTGATG 1447
465 pGlnSerAspTyrgLysValSerIleProIleGluGlyProTyrgLysProProH 482
1448 CTTACCGTTAT 1458
482 lstyArgTyr 485
seq_name: pir2:T06764
seq_documentation_block:
adenosylhomocysteine (EC 3.3.1.1) - wheat
N:Alternate names: S-adenosyl-L-homocysteine hydrolase
C:Species: Triticum aestivum (common wheat)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T06764
R:Richards, K.D.; Gardner, R.C.
A:Description: The influence of aluminium on histone, heat shock and S-adenosyl-L-hom
A:Residues: 1-485 <RIC>
A:Reference number: 215795
A:Accession: T06764
A:Status: translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Cross-references: EMBL:L11872; NID:g170772; PIDN:AAA34303.1; PID:g170773
C:Genetics:
A:Gene: SH6.2
C:Superfamily: adenosylhomocysteine
C:Keywords: NAD; thioether hydrolase
F:263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
F:86,120/Active site: Cys #status predicted
alignment_scores:
Quality: 1524.00 Length: 487
Ratio: 3.868 Gaps: 4
Percent Similarity: 80.903 Percent Identity: 61.807
```

alignment\_block:

US-09-759-990-1 x T06764

Align seg 1/1 to: T06764 from: 1 to: 485

```

19 ACTGTGCTCCATTCCGAGTACACAAATGCGGACATCAACCTCCATGTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluThrLysValLysAspLeuPheGlnAlaAspPh 24
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGCCGCTAAGGAACCTTACCTTCTGCTGAGAAAGAAATGCCAGGCTCTTGG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMeta 41
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTCCTTCGTCAGGCTTATTCCTGCTTCAAGCCATGAAGGGTGTGAGATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGTCTCCCTCCACATGACAGTCCAGACAGCGCTCCTCATCGAGACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTTGTGTGATGTACAGATGGGCTTCTGCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerS 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAGATACACCGCTGCTGCTATCGTTGCTCGGCCCAACAGGCACACCA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 erGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GAGAACCCAGCGGTATCCAGTCTTCGCTGGAAGGGCGCAACACTCCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTrpLysGlyGluThrLeuGl 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATPACTGGGAGAACATACATCCGCTCTCACATGCCAGTGGTCAAG 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 uGluTrpTrpTrpCysThrGluArgCysLeuAspTrpGlyValGlyGly 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 GCCACACAGCTGTCGATGATGTGGTGTGATGCTACATCTTCATCTCC 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 lProAspLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHis 148
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 AAGGCTTC.....CAATTCGAACAGCGGCTGTCTCCAGA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 GluGlyValLysAlaGluGluPheGluLysSerGlyLysValProAs 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
507 GCCAACAGAGCTGACACCTCGAATACCGCTGCGTCTTGTCTACACTCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 pProGluSerThrAspAsnProGluPheLysIleValLeuThrIleIleA 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
557 AGCAGGCTCTCAACCAACAGAACCACTGCGCACACAGTTGCTGCGGCG 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 rGaspGlyLeuLysThrAspAlaSerLysTrpArgLysMetLysGluArg 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
607 ATGAACCGTGTTCGAGAGACACACACAGGTGTCCACGCGCTCTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTrpGl 215
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 GCTCGAGNAGAGGCAACCTCTCTCCAGCCATCAACGTCACAGCAGC 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 nMetGlnGluSerGlyThrLeuLeuPheProAlaIleAsnValAsnAsps 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707 CTGTTTACAAGTCAAGTTCGATACATCTACGCTGCGCGCCACTCCCTT 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 erValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757 ATCGATGGTATCAACCGGCTTCGATGTCATGATCGCGCGCAAGACAGC 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807 TCTCGTATGGTTACGCGGATGTCGCGCAAGGGTGGCTCAATCCCTCC 856
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 aValValCysGlyTrpGlyAspValGlyLysGlyCysAlaAlaLeuL 282

```

```

857 GTGCCCAAGGCGCTGCTTATCATCAGAACTCGACCCCAATCTCGCT 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 ySlnAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
907 CTCAGGCTGCCATGGAAGGTACAGGTCCCGCGCATCGAGGAAGTCGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 LeuGlnAlaLeuMetGluGlyIleGlnIleLeuThrLeuGluAspValVa 315
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
957 CAAGGATGCGATATCTCGTTACATGACAGGAAACTCGGATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 lSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1007 CTGTTGATGATGCGCCAGATGAAGGATAAGGCTATTCTCGGTAACATC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 etValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 348
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1057 GGCACCTTCGATTAACGAATATGATACAGATGGCTCATGAAATACCCAG 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 GlyHisPheAspAsnGluIleAspMetAsnGlyLeuGluThrTyProGl 365
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1107 CATCAACGACATCCCAATCAAGCCAGATACACATGTGGGAATCCCAAG 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 yValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProG 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1157 ATGCCAC...GCTATCTCTCTTGTGAGGGCGCTTCTTAACTT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 luthrLysThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1204 GGTGCTGCTACAGTACCCCATCTTCTGTTATGTCAATGTCAATCACAAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1254 CCAGACACTCGCTCAGCTCGACCTCTAGAAAAGAGA.....GGAATC 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 nGlnValIleAlaGlnLeuGluLeuTrpAsnGluLysAlaSerGlyLysT 432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1298 TCAGAGAAGAGTTTACACACTTCCGAGCATCTCGATGAAGAAGTCGCT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 YrGluLysLysValTyValLeuProLysHisLeuAspGluLysValAla 448
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1348 CGCTCCACTCGGATCTCGATGTCCACCTTACACCTTACAAAGCTTACAGAA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449 AlaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLysSe 465
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1398 GCAGGCTGCTACTACATCAACGTTCAGTTGAGGTCCTTACAGTCTGATG 1447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465 rGlnSerAspTyIleSerIleProIleGluGlyProTyLysLeuArgL 1482
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1448 CTTACCGTTAT 1458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 euTyArgTyArg 485

```

seq\_name: pir2:S38379

seq\_documentation\_block:

N;Alternate names: S-adenosyl-L-homocysteine hydrolase

C;Species: Catharanthus roseus (Madagascar periwinkle)

C;Date: 22-Jan-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999

C;Accession: S38379

R;Schroeder, G.; Waitz, A.; Hotze, M.; Schroeder, J.

submitted to the EMBL Data Library, October 1993

A;Description: cDNA for stress-induced S-adenosyl-L-homocysteine hydrolase from Madag

A;Reference number: S38379

A;Accession: S38379

A;Molecule type: DNA

A;Residues: 1485 <SCH>

A;Cross-references: EMBL:Z26881; NID:g407411; PIDN:CAA81527.1; PID:g407412

C;Genetics:

C;Superfamily: adenosylhomocysteinase

C;Keywords: NAD; thioether hydrolase

F:263-292/Region: beta-alpha-beta NAD nucleotide-binding fold  
F:86,120/Active site: Cys #status predicted

alignment\_scores:  
Quality: 1520.00 Length: 488  
Ratio: 3.848 Gaps: 6  
Percent Similarity: 80.943 Percent Identity: 62.295

alignment\_block:  
US-09-759-990-1 x S38379 ..

Align seg 1/1 to: S38379 from: 1 to: 485

```

19 ACTGGTGTCTTCCATTCAGTACAGATTGCCGACATCAACCTCCATGTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGGCGCTAAGGAACCTTACCCTTGCCTGAGAGGAATGCCAGGTCTTTATGG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuMets 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTCCTGCTGAGGCTTATCCGCTTCTTAAGCCATTGAAGGTGTCAAGTC 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 exCysArgAlaGluPheGlyProSerGlnProPheLysGlyAlaLysIle 57
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGCTTCCCTCCATGACAGCTCCAGCAGCGCTCTCATCGAGACACT 218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTTGTGTGATGTGATGCTGAGTGGCTTCTGCAACATCTTCTTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTyrCysSerCysAsnIlePheSerT 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAGATACACCGCTGTCTGCTATGCTGTCGGCCCAACAGCCACACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 hrGlnGluHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GAGAAGCCAGCGGTATCCAGTCTTCGCTGCTGAGAGGGGAAACACTCCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTrpLysGlyGluThrLeuG1 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATACTGGGAGACACATACCGCCTCTCACATGG...CGAGATGGTC 415
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 nGluTyrTrpTyrCysThrGluArgAlaLeuAspTrpGlyProAspGly. 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 AAGGCCACAGCAGGTGTGATGATGGTGGTGTGATGTACACTCCTCATC 465
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 ..GlyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeulle 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466 TCCAGGGGCTC.....GAATTCGAACACAGCGGTGCTGTGTC 503
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 HisGluGlyValLysAlaGluGluTyrLysLysAsnGlyAlaLeuPr 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 AGAGCCACAGAGCTGACACCTCCGAATACCGCTGGTCTTCTGCTACAC 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 oAspProSerSerThrAspAsnAlaGluPheGlnIleValLeuThrIle 181
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
554 TCAGCAGGTCTTCAACCAAGACAAAGAACACCTGGCACACAGTTGCTGCC 603
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 leArgAspGlyLeuLysSerAspProThrLysTyrThrArgMetLysGlu 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
604 GGATGAAGCGGTCTTCCGAGAGACAAACAGAGTGTCTCCACCGCTCTTA 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 ArgLeuValGlyValSerGluGluThrThrGlyValLysArgLeuTy 214
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
654 CCAGCTCGAAGAGGGGAACTCTCTTCCAGCCATCAACGTCAACG 703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 rGlnMetGlnAlaAsnGlyThrLeuLeuPheProAlaIleAsnValAsnA 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
704 ACAGCTGTACAAAGTCCAAAGTTCGATAACATCTACGGCTCCCGCCACTCC 753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 spSerValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSer 247

```

```

754 CTTATCGATGGTATCAACCGTGTTCGATGTCATGATCGCGGCAAGAC 803
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
248 LeuProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVa 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
804 AGTCTCGTCATGGGTTACGGCGATGTCGGCAAGGGTGCCTCAATCCC 853
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 lAlaValAlaGlyTyrGlyAspValGlyLysGlyCysAlaAlaL 281
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
854 TCCGTGCCCAAGGCGTCCGTTTATCATCACAGAACTCGACCCCAATCGC 903
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 euLysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCys 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
904 GCTCTCCAGGCTGCCATGGAAGGTACCAGTCCGCCGCATCGAGGAAGT 953
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298 AlaLeuGlnAlaThrMetGluGlyLeuGlnValLeuThrLeuGluAspVa 314
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
954 CGTCAAGGATGTCATATCTTCGTTACATGCACAGAACTCGCATATCA 1003
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 lValSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleI 331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1004 TCTCTGTTGACATGATGCCCGCAGATGAAGGATAAGGCTATTGTCGGTAAC 1053
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331 leMetValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsn 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1054 ATCGGCCACTTCGATAAGAAATGATGATGATGATGATGATGATGATGAT 1103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
348 lIleGlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyPr 364
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1104 AGCATCAAGCACATCCCAATCAAGCCAGATACGACATGTGGGAATTCC 1153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 oGlyValLysArgIleThrIleLysProGlnThrAspArgTrpValPheP 381
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1154 CAGATGCCACGCT...ATCCTCTCTTCTGTGAGGCGCGCTTCTTAAC 1200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
381 roAspThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsn 397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201 CTTGGTCCGCTACAGTACCCCATCTTTCGTTATGTCATGTCATTCAC 1250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 LeuGlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheTh 414
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1251 AAACAGACACTCGCTCAGCTCAGCTCTACGAAAGAGA.....GGAA 1294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 rAsnGlnValIleAlaGlnLeuGluLeuTrpAsnGluArgLysThrGlyL 431
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1295 ATCTCGAAGAAGGTTTACACACTTCGAGCACTCTCGATGAAGAATC 1344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 ystyrGluLysLysValTyrValLeuProLysHisLeuAspGluLysVal 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1345 GCTCGCTCCACCTCGATCTCTCGATGTCCACCTTACAAAGCTTACACA 1394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 AlaAlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThrly 464
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1395 GAAGCAGGCTGACTACATCACTTCCAGTTCAGGGTCTTACAAAGTCTG 1444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
464 sAspGlnAlaAspTyrIleSerValProIleGluGlyProTyrLysProA 481
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1445 ATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 laHisTyrArgTyr 485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: pir2:S71621

```

seq\_documentation\_block:  
adenosylhomocysteinase (EC 3.3.1.1) - Phalaenopsis sp.  
N;Alternate names: S-adenosyl-L-homocysteine hydrolase  
C;Species: Phalaenopsis sp.  
C;Date: 27-Nov-1997 #sequence\_revision 12-Dec-1997 #text\_change 22-Jun-1999  
C;Accession: S71621  
R;Preisig-Mueller, R.; Gnaul, P.; Kindl, H.  
Arch. Biochem. Biophys. 317, 201-207, 1995  
A;Title: The inducible 9,10-dihydrophenanthrene pathway: characterization and express  
A;Reference number: S71619; MUID:95177653

A:Accession: S71621  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-485 <PRE>  
 A:Cross-references: EMBL:X79905; NID:9758246; PIDN:CAA56278.1; PID:g758247  
 C:Experimental source: clone SH511  
 C:Superfamily: adenosylhomocysteinase  
 F:Keywords: NAD; thioether hydrolase  
 F:263-292/Region: beta-alpha-beta NAD nucleotide-binding fold  
 F:86,120/Active site: Cys #status predicted

alignment\_scores:  
 Quality: 1509.00 Length: 482  
 Ratio: 3.889 Gaps: 4  
 Percent Similarity: 80.498 Percent Identity: 62.656  
 alignment\_block:  
 US-09-759-990-1 x S71621 ..  
 Align seg 1/1 to: S71621 from: 1 to: 485

```

34 GAGTACAGAAATGCGGACATCAACCTCCATGTTCTCGCGCGTAAGGAAC 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GluTyrLysValLysAspLeuSerGlnAlaAspPheGlyArgLeuGluI 29
84 TACCTTGCTGAGAAAGGAATGCCAGGTCTTATGTTCTTCCTGAGCGTT 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgAlaGlu 46
134 ATTCGGCTTCTAAGCCATTGAAGGTGTGCAATCTCTGTTCCCTCCAC 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 heGlyProSerGlnProPheLysGlyAlaArgLysSerLeuHis 62
184 ATGACATGTCAGACAGCGCTCCTCATCGAGACTCAAGCTCTTGGTGC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 79
234 TGATGCAGATGGCTTCCTGCAACATCTTCTTACACAAGATACAGCG 283
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 aGluValArgTipCysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTCTATGTTGTCGGCCCAACAGGCGACACAGAGAACCCAGCGGT 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 laAlaAlaIleAlaArgAspSerAlaAla..... 105
334 ATCCCAAGTCTTCGCGGAGGCGGAACACTCCAGAACTACTGGGAGAA 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTrpLysGlyGluThrLeuGlnGluTyrTrpCy 120
384 CACATACCGCGTCTCACATGCGCAGATGTCAGGCCACAGCGGTG 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 sThrGluArgCysLeuGluTrpGlyAlaGlyGlyGlyProAspLeuIle 137
434 TGGATGATGGTGTGATGATGATGATGATGATGATGATGATGATGAT 477
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 alAspAspGlyGlyAspAlaThrLeuLeuIleHisGlyValLysAla 153
478 .....GAATTCGAACAGCGGTGCTGCTTCCAGAGCCACAGAGCTGA 521
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 GluGluGluTyrGluLysAsnGlyLysIleProAspProAlaSerThr 170
522 CAACCTCGAATACCGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 pAsnAlaGluPheGlnIleValLeuGlyLeuIleArgAspSerLeuSer 187
572 AAGACAAAGAACCACTGGCACAGATGCTGCTGCGGCATGACGGTGTTC 621
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 alAspProLysLysTyrArgArgMetLysGluArgLeuValGlyValSer 203
622 GAAGAGACAAACAGGTGTCCACCGCTCTTACCAGCTCGAGAGGAGGG 671
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 GluGluThrThrThrGlyValLysArgLeuTyrGlnMetGlnTyrSerGI 220

```

```

672 CAAACTCCTCTTCCACGCCATCAAGTCAACGACGCTGTTACAAAGTCCA 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 YThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSerL 237
722 AGTTCGATAACATCTACGGCTGCGCGCACCTCCCTTATCGATGGTATCA 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 yspPheAspAsnLeuTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253
772 CGTGTTCGATGTCATGATCGGGCGGCAAGACAGCTCTCTGTCATGGGTTA 821
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValValCysGlyTy 270
822 CGCGGATGTCGCAAGGCTGCGCTCAATCCCTCCGTCGCGCAAGCGGTC 871
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 rGlyAspValGlyLeuGlyCysAlaAlaLeuLeuLysThrAlaGlyAla 287
872 GCGTTATCATCAGAACTCGACCCCAATCTGCGCTCTCCAGGTCGCCATG 921
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 rgValIleValThrGluIleAspProIleCysAlaLeuGlnAlaLeuMet 303
922 GAAGCTACCGGTCGCGCATCGAGGAAGTCTCAAGGATGTCGATAT 971
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 GluGlyLeuProValLeuArgLeuGluAspValValSerGluAlaAspI 320
972 CTTCTGTTACATGCACAGGAACCTCGCATATCATCTCTGTTGACATGATG 1021
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ePheValThrThrGlyAsnLysAspIleIleMetValAspHisMetA 337
1022 CCCAGATGAAGGATAAGGCTATTGTCGTAACATCGGCCACTTCGATAAC 1071
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 rgLysMetLysAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353
1072 GAAATTGATACAGATGGCTCATGAATACCCAGGATCAAGCACATCCC 1121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 GluIleAspMetLeuGlyLeuGluSerPheProGlyValLysArgIleTh 370
1122 AATCAAGCCAGAATACGACATGTGGAAATTCGAGATGGCCAGCT...A 1168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 rIleLysProGlnThrAspArgValPheProAspThrAsnSerGlyI 387
1169 TCCCTCTCTTCTGTCAGGCGCCCTTCACTTAACTTGGCTGCGGTACAGGT 1218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 leuValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
1219 CACCATCTTTCGTTATGTCATGTCATTACAAACCCACAGACTCGCTCA 1268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 HisProSerPheValMetSerSerSerPheThrAsnGlnValIleAlaGI 420
1269 GCTCGACCTCTACGAAAGAGA.....GGAAATCTCGAGAAAGGTTT 1312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 nLeuGluLeuTrpLysGluArgAlaSerGlyLysTyrGluLysLysVal 437
1313 ACACACTTCGGAACATCTCCATGAAGAAGTCGCTGCGCTCCACCTCGGA 1362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
437 yrValLeuProLysHisLeuAspGluLysValAlaAlaLeuHisLeuGly 453
1363 TCCTCGATGTCACCTTACAAAGCTTACAGAGAGGCTGACTACAT 1412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 LysLeuGlyAlaLysLeuThrLysLeuThrProSerGlnAlaAspTyrI 470
1413 CAACGTTCCAGTTCAGGTCCTTCAAGCTCTGATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 eSerValProValGluGlyProTyrLysProAlaHisTyrArgTyr 485

```

seq\_name: pir2:D82730

seq\_documentation\_block:

adenosylhomocysteinase XF1037 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: D82730  
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

125 lySerThrTrpValAspGluPro.....SerAspSerLeuGluGluGln 139  
538 TCGGTCCTGCTACACTCAAGCAGGCTCTTCAACCAAGACAAAGAACCACTG 587  
140 ValIleArgLeuLeuLysArgIleAlaIleGluArgProGlyTyrTr 156  
588 GCACACAGTTGCTCGCGCATGAACGGTGTTCGGAAGAGACAAACACAG 637  
156 pThrArgValValAsnAspTrpLysGlyValSerGluGluThrThr 173  
638 GTGTCACCGCCCTCTACAGCTCAGAGGAGGAGGAGGAGGAGGAGGAGG 687  
173 lyValHisArgLeuTrpGlnIleAlaAlaThrGlyArgLeuLeuValPro 189  
688 GCATCAACAGCTCAACAGCGCTGTTCACAAAGTCCAAAGTTCGATACTCTA 737  
190 AlaIleAsnValAsnAspSerValThrLysSerLysPheAspAsnLeuTy 206  
738 CGGTCGCGCCACTCCCTTATCGATGATGATCAACCGTGTTCGATGTC 787  
206 rGlyCysArgGluSerLeuAlaAspGlyLeuLysArgAlaMetAspValm 223  
788 TGATCGCGCGCAAGACAGCTCTCGTCATGGGTACGGCGATGTCGGCAAG 837  
223 etLeuAlaGlyLysLeuAlaValValCysGlyTyrGlyAspValGlyLys 239  
838 GGTCGCGCTCAATCCCTCGGTGGCGCAAGCGCTCGGTATTCATCATCACA 887  
240 GlySerAlaHisSerLeuArgAlaTyrGlyAlaArgValIleValThr 256  
888 ACTCGACCCATCTCGCTCTCCAGGCTGCGATGGAAGCTTACAGAGTCC 937  
256 uileAspProIleCysAlaLeuGlnAlaAlaMetGluGlyPheGluValT 273  
938 GCCGATCGAGGAAGTCTCAAGGATGTCGATATCTTCGTTCATATGCA 987  
273 hrThrValGluAspThrLeuGlyGlnAlaAspIleTyrValThrThr 289  
988 GGAACATCGGATATCACTCTGTTGCACATGATGCGCCAGATGAAGATAA 1037  
290 GlyAsnLysAspValIleArgIleGluHisMetThrAlaMetLysAspGl 306  
1038 GGCTATTGCGGTAACATCGCGCTTCGATTAACGAAATGTATACAGATG 1087  
306 nValIleValCysAsnIleGlyHisPheAspAsnGluIleGlnValAsp 323  
1088 GCTCATGAATACCCAGGATCAAGCATCCCAATCAAGCCAGCAATAC 1137  
323 laLeuAsnAlaLeuAlaGlyValGlnLysIleAsnIleLysProGlnVal 339  
1138 GACATGTCGGAATTCACAGATGCGCCAGCTATCTCTCTTCTGCTGAGG 1187  
340 AspLysPheIleLeuProAsnGlyAsnThrLeuPheLeuAlaGluGl 356  
1188 CGGCTCTCTTAACCTGCGCTACAGGTACAGGTACCCATCTTCGTTATGT 1237  
356 yArgLeuValAsnLeuGlyCysAlaThrGlyHisProSerPheValMet 373  
1238 CAATGTCATTCACAAACAGGCTTTCACACATTCGGAAGCATCTCGATGA 1287  
373 erAsnSerPheAlaAsnGlnThrLeuAlaGlnIleAspLeuTrpGlnAsn 389  
1288 AGAGGAAATCTCGAGAAGAGGTTTACACATTCGGAAGCATCTCGATGA 1337  
390 LysAspValTyrGluLysAsnValTyrArgLeuProLysLysLeuAspGl 406  
1338 AGAAGTCGCTCGCTCCCTCCAGTCTCTCGATGTCCTCCATCTTACAAAGC 1387  
406 uGluValAlaArgLeuHisLeuGluLysIleGlyValLysLeuThrThr 423  
1388 TTACACAGAGGAGGCTGATCATCATCAACGTTCCAGTTGAGGTCCTTAC 1437  
423 euThrAlaAsnGlnAlaAlaTyrLeuGlyIleSerValGluGlyProPhe 439

Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: D82730  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-446 <SIM>  
A:Cross-references: GB:AE003941; GB:AE003849; NID:99105978; PIDN:AAF83847.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H  
as-Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.F.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XFL037  
C:Superfamily: adenosylhomocysteinase  
alignment\_scores:  
Quality: 1406.00 Length: 457  
Ratio: 3.790 Gaps: 5  
Percent Similarity: 81.182 Percent Identity: 60.613  
alignment\_block:  
US-09-759-990-1 x D82730 ..  
Align seg 1/1 to: D82730 from: 1 to: 446  
103 ATGCCAGGCTTATGGTCTTCTCGTAGCGGTATTCGCGTCTTAAGCCATT 152  
1 MetProGlyLeuMetSerIleArgArgLysTyrAlaSerLysGlnProLe 17  
153 GAAGGGTGTGAGAATCTCGTTTCCCTCCACATGACAGTCCAGACAGCGG 202  
17 uLysGlyValArgValThrGlySerLeuHisMetThrIleGlnThrAlaV 34  
203 TCCATATCAGACACTCAGACTCTTGGTGTCTGCTGCTGCTGCTGCTGCTTC 252  
34 alteuileGluThrLeuLysAspIleGlyAlaAspValArgTrpAlaSer 50  
253 TGCACATCTTCTACACAAGATACAGCGCTGCTGCTGCTGCTGCTGCTGCG 302  
51 CysAsnIlePheSerThrGlnAspHisAlaAlaAlaIleAlaThr.. 66  
303 CCCAACAGGCACACAGAGAACGCCGGTATCCCGAGTCTTCGCTCGGA 352  
67 .....SerGlyThrProValPheAlaTrpL 75  
353 AGCGGNAACACTCCAGAGTACTGGGAACACATACCGCGCTCTCACCA 402  
75 yGlyGluThrLeuGluGluTyrTrpAspCysThrLeuGlnAlaLeuThr 91  
403 TGG.....CCAGATGGT.....CAAGGCCACAGCAGGTGTTCGATGA 440  
92 PheThrLeuSerAspGlyThrLeuThrGlyProGluLeuIleValAspAs 108  
441 TGGTGTGATGCTACACTCTCTCTCCAGGCTTCCAAAGGCTTCCAAATCGAAACAG 490  
108 pGlyGlyAspAlaThrLeuLeuIleHisLysGlyTyrGluLeuGluAsnG 125  
491 CCGGTGCT...GTTCCAGAGCCACAGAGAGTGCACACCTCGAATACGCG 537  
..... ||| ||||| :||||| :|||



1438 AGTCTGATGCTTACGGTTAT 1458  
 ||| :|||  
 440 LysProGluHisTyrArgTyr 446

seq\_name: pir2:D97362

seq\_documentation\_block:

adenosylhomocysteine (S-adenosyl-L-homocysteine hydrolase) (adocytase) [imported] - Ag  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
 C:Accession: D97362

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; PMID:11743194

A:Accession: D97362

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-466 <KOR>

A:Cross-references: GB:AE007869; PIDN:AAK85853.1; PID:gi1514892; GSPDB:GN00169

C:Genetics:

A:Gene: AGR\_C\_46

A:Map position: circular chromosome

C:Superfamily: adenosylhomocysteinease

alignment\_scores:

Quality: 1382.00 Length: 478  
 Ratio: 3.627 Gaps: 6  
 Percent Similarity: 79.707 Percent Identity: 58.577

alignment\_block:

US-09-759-990-1 x D97362 ..

Align seg 1/1 to: D97362 from: 1 to: 466

34 GAGTACAGAAATGCGGACATCAACCTCCATGTCGCGCGGTAAAGAACT 83  
 ||| :|||  
 6 AspTyrIleValAlaAspIleAsnLeuAlaPheGlyArgLysGluLe 22  
 84 TACCTTGCTGAGAGGAAATGCCAGGCTTATGTTCTCTGCTGAGCGTT 133  
 | :|||  
 22 uAspIleAlaGluThrGluMetProGlyLeuMetSerCysArgLysGluP 39  
 134 ATTCGGCTTTAAGCATTAAGGGTGTGAGAAATCTCTGTTCCCTCCAC 183  
 :|||  
 39 heGlyGluSerLysProLeuLysGlyAlaArgIleThrGlySerLeuHis 55  
 184 ATGACAGTCCAGACAGCGGCTCTCATCGAGACACTCACAGCTCTGGTGC 233  
 ||| :|||  
 56 MetThrIleGlnThrGlyValIleuIleGluThrLeuLysGluLeuGlyAl 72  
 234 TGATGTCAGATGGCTTCTCTCAACATCTCTCTACAGAGATACAGCGG 283  
 :||| :|||  
 72 aGluIleArgTyrAlaSerCysAsnIlePheSerThrGlnAspHisAla 89  
 284 CTGCTGCTATCTGTTGTCGGCCCAACAGGACACACAGAGAGCCGCGGT 333  
 ||| :|||  
 89 laAlaIleAlaAla.....AlaGly 96  
 334 ATCCAGCTCTTGGCTGGAAGGGGGAACACTCCAGCATCTACTGGGAA 383  
 ||| :|||  
 97 IleProValPheAlaValLysGlyCysLeuSerLeuThrGluTyrTrpGlu 113  
 384 CACATACCGGCTCTCATATGCCAGATGGTCAAGGCCACACAGAGGTTG 433  
 ||| :|||  
 113 rThrAspLysIlePheGlnThrTrpAspGlyLeuSerAsnMetIleL 130  
 434 TCATGATGGTGGTGTACATCTCTCATCTCTCAAGGGCTTCGAAATTC 483  
 :||| :|||  
 130 euAspAspGlyGlyAspAlaThrMetTyrIleLeuLeuGlyAlaArgAla 146

484 GAAACAGCC...GGTGTCTGTTCCAGAGCCA...ACAGAAGCTGACAACCT 527  
 ||| :|||  
 147 GluAlaGlyGluaspValLeuSerAsnProGlySerGluGluGluGlu 163  
 528 CGAATACCGCTGGCTTCTTCTACACTCAAGCAGGTCTTCAACCAAGACA 577  
 :||| :|||  
 163 eLeuPheAlaGlnIleAsnLysArgLeuLysAla.....S 175  
 578 AGAACCACTGGCACACAGTGTCTGCC...GGCATGAACGGTGTTCGGA 624  
 :||| :|||  
 175 erProGlyTrpPheThrLysGlnArgAspAlaLeuLysGlyValThrGlu 191  
 625 GAGACAACAACAGGTGTCCACCGCTCTTACCAGCTCGAGAGAGGGGCA 674  
 ||| :|||  
 192 GluThrThrThrGlyValHisArgLeuThrAspLeuAlaLysGlyLe 208  
 675 ACTCTCTTCCAGCCCATCAACGCTCAACGAGCTGTGTACAAAGTCCAGT 724  
 ||| :|||  
 208 uLeuProPheProAlaIleAsnValAsnAspSerValThrLysSerLysP 225  
 725 TCGATAACATCTACGGCTCGCGCACCTCCCTTATCGATGGTATCAACGT 774  
 ||| :|||  
 225 heAspAsnLysTyrGlyCysLysGluSerLeuValAspGlyIleArg 241  
 775 GCTTCCGATGTCTATGATCGCGGCAAGACAGCTCTCGTCATGGGTACGG 824  
 ||| :|||  
 242 AlaThrAspValMetMetAlaGlyLysValAlaValValCysGlyTyrGl 258  
 825 CGATGTGCGCAAGGGTGTGCTCAATCCCTCGGCGCAAGGGCTCGCG 874  
 ||| :|||  
 258 yAspValGlyLysGlySerAlaAlaSerLeuGlnGlyAlaGlyAlaArg 275  
 875 TTATCATCACAGAACTCGACCAATCTCGCTCTCCAGGCTGCCATGAA 924  
 ||| :|||  
 275 aLysValThrGluIleAspProIleCysAlaLeuGlnAlaAlaMetAsp 291  
 925 GCTTACAGGTCCGCGCATCGAGGAAGTGTCAAGGATGTGCGATATCTT 974  
 ||| :|||  
 292 GlyPheGluValValArgLeuGluAspValValSerSerAlaAspIlePh 308  
 975 CTTTACATGCACAGGAACTGCGATATCTCTGTGTGACATGATGGCC 1024  
 ||| :|||  
 308 eIleThrThrThrGlyAsnLysAspValIleArgIleGluHisMetArg 325  
 1025 AGATGAAGGATAAGGCTATTGTGCGTAACATCGCCACTTCGATAACGAA 1074  
 :||| :|||  
 325 luMetLysAspMetAlaIleValGlyAsnIleGlyHisPheAspAsnGlu 341  
 1075 ATTGATACAGATGGCTCATGAAATACCCAGGCATCAAGCACATCCCAAT 1124  
 ||| :|||  
 342 IleGlnValAlaSerLeuArg.....AsnLeuLysTrpThrAsnI 355  
 1125 CAAGCCAGAAATACGATGTTGGGAATCCAGATGGCCAGCTATCTCTCC 1174  
 ||| :|||  
 355 eLysProGlnValAspMetIleGluPheProLysGlyAsnArgIleLeL 372  
 1175 TCTTGTCTGAGGCGCGCTTTCTTAACCTTGGCTGCGCTACAGGTCAACCA 1224  
 ||| :|||  
 372 euLeuSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisPro 388  
 1225 TCTTGTGTATGTCATGTCATTCACAAACCCAGACTCGCTCAGCTCGA 1274  
 ||| :|||  
 389 SerPheValMetSerAlaSerPheThrAsnGlnValLeuGlyGlnIleGl 405  
 1275 CCTCTACGAAAGAGAGGAAATCTCGAGAGAGAGGTTTACACACTTCGA 1324  
 :||| :|||  
 405 uLeuPheThrLysProGlyGlyTrpLysAsnGluValThrValLeuProL 422  
 1325 AGCATCTCGATGAAGAAGTCGCTCGCTCCACCTCGGATCTCTCGATGTC 1374  
 ||| :|||  
 422 yHisLeuAspGluLysValAlaArgLeuHisLeuGluLysLeuGlyVal 438



```

405 uLeuPheThrLysProGlyLysAsnGluValTyValLeuProL 422
1325 AGCATCTCGATCAAGAGTCGCTCCGCTCCACCTCGCATCTCTCGATGTC 1374
422 yHisLeuAspGluLysValAlaArgLeuHisLeuGluLysLeuGlyVal 438
1375 CACCTTACAAAGCTTACACAGAGGAGGCTGACTACATCAACGTTCCAGT 1424
439 ArgLeuThrGluLeuSerAspLeuGlnAlaAspTyrIleGlyIleSerLy 455
1425 TGAGGCTCTTACAGTCTGATGCTTACCGTTAT 1458
455 sGlnGlyProPheLysAlaGluHisTyrArgTyr 466
seq_name: pir2:H87280

```

```

seq_documentation_block:
adenosylhomocysteinase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: H87280
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173696; PMID:11259647
A:Accession: H87280
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <STO>
A:Cross-references: GB:AE005673; NID:g13421390; PIDN:AAK22244.1; GSPDB:GN00148
C:Genetics:
C:Gene: CC0257
C:Superfamily: adenosylhomocysteinase

```

```

alignment_scores:
Quality: 1374.00 Length: 475
Ratio: 3.606 Gaps: 3
Percent Similarity: 80.211 Percent Identity: 58.105
alignment_block:
US-09-759-990-1 x H87280
Align seg 1/1 to: H87280 from: 1 to: 463

```

```

34 GAGTACAGAAATGCCGACATCAACCTCCATCTTCGGCCGTAAGAACT 83
3 AspTyrIleValLysAspIleSerLeuAlaAspPheGlyArgLysGluI 19
84 TACCCTTGCTGAGAGGAATGCCAGGTCTTATGTTCTTCGTGAGCGTT 133
19 eAlaIleAlaGluThrGluMetProGlyLeuMetAlaThrArgAlaGlu 36
134 ATTCGCTTCTAAGCCATTTGAAGGTGTGAGAACTCTGTTCTCCCTCCAC 183
36 yrGlyProGlnGlnIleLeuLysGlyAlaArgIleAlaGlySerLeuHis 52
184 ATGACAGTCAGACAGCGCTCTCATCGACACTCACAGCTCTTGGTGC 233
53 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 69
234 TGATGTCAGATGGCTTCCTGCAACATCTTCTTACACAGATACAGCG 283
69 aGluValArgThrPalaSerCysAsnIlePheSerThrGlnAspHisAla 86
284 CTGCTGCTATCTGTTGTCGGCCCAACAGGCCACACCAGAGAACCCAGCGGT 333
86 laAlaAlaIleAlaAla.....AlaGly 93
334 ATCCAGTCTTCGCTGGAGGCGGAACACTCCCAAGATACTGGGAGAA 383

```

```

94 ValProValPheAlaPheLysGlyGluAsnLeuValGluTyrTrpGluTy 110
384 CACATACCCGGCTCTCATGCGCAGATGGTCAAGGCCACAGCAGGTG 433
110 rAlaHisLysIlePheGluTrpHisAspGlyGlyTyrProAsnLeuIleL 127
434 TCGATGATGTTGTTGATGCTACACTCTCATCTCCAAAGGCTTCGAATTC 483
127 euAspAspGlyGlyAspAlaThrLeuLeuCysValLeuGlyProLysAla 143
484 GAAACAGCCGGTCTCTTCAGAGCCCAACAGAGCTGACACCTCGAATA 533
144 GluLysAspProSerIle.....LeuAsnAsnProGlnAsnGluGluG 158
534 CGGCTCGCTTCTGCTACACTCAAGCAGGTCTTCAACCAAGCAAGAAC 583
158 uGluAlaLeuTyrAlaValMetLysLysTyrLeuAlaGluLysProGly 175
584 ACTGGCACACAGTTGCTGCGGCATGAACGGTGTTCGGAAGACAGACA 633
175 heTy-SerAlaIleArgAlaAlaIleGlyGlyValSerGluGluThr 191
634 ACAGTGTCCACCGCTCTACAGCTCGAGAGAGGAGGCAACTCTCTT 683
192 ThrGlyValHisArgLeuTyrGlnMetAlaGlnLysAspGluLeuPro 208
684 CCACGCCATCAACGTCACAGCCTGTTACAAAGTCCAAAGTTCGATACA 733
208 eProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspAsn 225
734 TCTACGGCTCGCCCACTCCCTTATCGATGGTATCAACCGTCTCGAT 783
225 euTyrGlyCysArgGluSerLeuValAspAlaIleArgArgGlyThrAsp 241
784 GTCATGATCGCGCGGCAAGACAGCTCTCGTCATGGTTACGGCGATGCG 833
242 ValMetLeuSerGlyLysValAlaValValCysGlyTyrGlyAspValG 258
834 CAAGGCTCGCTCAATCCCTCGTGGCCAAAGCGCTCGCGTTCATCATCA 883
258 yLysGlySerAlaAlaSerLeuArgGlnGlyAlaArgValIleVal 275
884 CAGAACTCGACCAATCTCGCTCTCCAGGCTGCCATGGAAGGCTACAG 933
275 hrGluValAspProIleCysAlaLeuGlnAlaAlaMetGluGlyTyrGlu 291
934 GTCCGCGCATCGAGGAAGTCGTCAGGATGTCGATATCTTCGTTACATG 983
292 ValGlnThrLeuAsnAspValAlaAspLysAlaAspIlePheValThrAl 308
984 CACAGGAACTCGCATATCATCTCTGTTGACATGATGCCCGCATGAAGG 1033
308 aThrGlyAsnLysAspValIleThrValAspAspMetArgLysMetLys 325
1034 ATAGGCTATTGTCGGTAACTGCGCCACTTCGATAACGAAATTTGATACA 1083
325 snAsnAlaIleValCysAsnIleGlyHisPheAspSerGluIleGlnIle 341
1084 GATGGCTCATGAATACCCAGGCATCAGACACATCCCAATCAAGCCAGA 1133
342 AlaGlyLeuArgAsnPhe.....LysTrpAspGluIleLysProG 355
1134 ATAGCAGATGTGGGAATTCAGAGATGGCCAGCTATCTCTCTCTGCTG 1183
355 nValHisHisValGluPheProAspGlyLysLysLeuIleValLeuSer 372
1184 AGGCCGCTCTTAACTTGGCTGCGGTACAGGTACACCATCTTTCTGTT 1233
372 luGlyArgLeuValAsnLeuGlyAsnAlaThrGlyHisProSerPheVal 388
1234 ATGTCAATGTCTTCAACAAACAGACACTCGCTCAGCTCGACCTCTACGA 1283

```

389 MetSerAlaSerPheThrAsnGlnThrLeuAlaGlnIleGluLeuTprTh 405  
1284 AAAGAGAGAAATCTCGAAGAGAGAGTTTACACACTCCGAGCATCTCG 1333  
405 rAsnLysAlaLysTyrGluAsnGlnValTyrThrLeuProLysHisLeuA 422  
1334 ATGAAGAGTGCCTCGCTCCACCTCGGATCTCTCGATGTCCACCTTACA 1383  
422 spGluLysValAlaPheLeuHisLeuGluLysLeuGlyAlaLysLeuThr 438  
1384 AAGCTTACACAGAGAGGCTGATCATCAATCAACGTTCCAGTTGAGGTC 1433  
439 ThrLeuArgLysAspGlnAlaAspTyrIleGlyValProGluAlaGlyPr 455  
1434 TTACAAGTCTCATGCTTACCGTTAT 1458  
455 oPheLysProAspHisTyrArgTyr 463  
seq\_name: pir1:A46035  
seq\_documentation\_block:  
adenosylhomocysteine (EC 3.3.1.1) - Rhodobacter capsulatus  
N; Alternate names: S-adenosyl-L-homocysteine hydrolase  
C; Species: Rhodobacter capsulatus  
C; Date: 04-Mar-1994 #sequence\_revision 04-Oct-1996 #text\_change 18-Jun-1999  
C; Accession: A46035; A36863  
R; Sganga, M.W.; Aksmit, R.R.; Cantoni, G.L.; Bauer, C.E.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6328-6332, 1992  
A; Title: Mutational and nucleotide sequence analysis of S-adenosyl-L-homocysteine hydrolase  
A; Reference number: A46035; MUID:92335291  
A; Accession: A46035  
A; Molecule type: DNA  
A; Residues: 1-462 <SGA>  
A; Note: sequence extracted from NCBI backbone (NCBIN:108691, NCBI:P:108696)  
R; Bugg, J.J.; Sganga, M.W.; Bauer, C.E.  
J. Bacteriol. 176, 61-69, 1994  
A; Title: Nucleotide sequence and characterization of the Rhodobacter capsulatus hvrB gene  
A; Reference number: A36863; MUID:94110241  
A; Accession: A36863  
A; Molecule type: DNA  
A; Residues: 1-13 <BUG>  
A; Cross-references: GB:L23836; NID:g577613; PIDN:AAA53540.1; PID:g577614  
C; Genetics:  
A; Gene: ahcy  
C; Function:  
C; Superfamily: adenosylhomocysteinease  
C; Keywords: NAD; thioether hydrolase  
F; 247-276/Region: beta-alpha-beta NAD nucleotide-binding fold

alignment\_scores:  
Quality: 1352.50 Length: 484  
Ratio: 3.636 Gaps: 9  
Percent Similarity: 76.860 Percent Identity: 59.298  
alignment\_block:  
US-09-759-990-1 x A46035 ..  
Align seg 1/1 to: A46035 from: 1 to: 462  
34 GAGTACAGAAATTCGACATCAACCTCCATGTCTCGCGGTAAAGACT 83  
3 AspTyrIleValLysAspIleLysLeuAlaGluPheGlyArgLysGluLe 19  
84 TACCCTGTCTGAGAGAAATGCCAGGTCTTATGTTCTTCGTGAGCGTT 133  
19 uAspIleAlaGluThrGluMetProGlyLeuMetAlaCysArgGluGluP 36  
134 ATTCCGCTTCTAAGCATTGAGGGTGTGAGATCTCTGAGATCTCTCCAC 183  
36 heGlyProSerGlnProLeuLysGlyAlaArgIleAlaGlySerLeuHis 52  
184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACACTCACAGCTCTTGTGTC 233  
53 MetThrIleGlnThrAlaValLeuIleGluThrLeuLysAlaLeuGlyAl 69  
234 TGAATGCATGCGGCTCTCGAAGAGAGTTTACACACTCCGAGCATCTCG 283  
69 aspValArgTyrPheLysSerCysAsnIlePheSerThrGlnAspHisAla 86  
284 CTGCTCTCATGCTGTGCGGCCAACACAGGCACACAGAGAGAGCCAGCG 333  
86 laAlaAlaIleAlaAlaGly.....Gly 93  
334 ATCCAGTCTTCCGCTGGAAGAGGGAACACTCCAGAAATCTGGAGAA 383  
94 ThrProValPheAlaValLysGlyGluThrLeuGluGluTyrTrpAlaTy 110  
384 CACATACCGGCTCTCATGTCAGGTCAGATGTCAGAGGCCACAGCAGGTTG 433  
110 rThrAspLysIlePheGlnPheProGluGlyThr....CysAsnMetIleL 126  
434 TCGATGATGCTGTGATGTACACTCTCTCATCTCCAAAGGCTTCGAATTC 483  
126 euAspAspGlyGlyAspAlaThrLeuTyrIleLeuLeuGlyAlaArgVal 142  
484 GAAACA.....GCCGCTGTGTTCAGAGCCCAACAGAGCTGACAACT 527  
143 GluAlaGlyGluThrAspLeuIleAlaThrProThrSerGluAspGluVa 159  
528 CGAATACCGGCTGCTTCTGCTACACTCAAGCAGGTCTTCAACCAAGACA 577  
159 l.....Cys.....LeuPheAsnGlnIleL 166  
578 AGAACAC.....TGGCACACA....GTTGCTGCGCGC 606  
166 yLysArgMetValGluSerProGlyThrPheThrGlnArgAlaAla 182  
607 ATGACGCTGTTCCGAGAGACACAAACAGGTGTCCACCGCTCTTACCA 656  
183 IleLysGlyValSerGluGluThrThrGlyValHisArgLeuTyrAs 199  
657 GCTCGAAGAGGAGGCAAACTCTCTTCCAGCCATCAACGTCACAGCAG 706  
199 pLeuHisLysLysGlyLeuLeuProPheProAlaIleAsnValAsnAsp 216  
707 CTGTTACAAAGTCCAACTTCGATACATCTACGCTGCGCGCTCCCTTT 756  
216 erValThrLysSerLysPheAspAsnLysTyrGlyCysLysGluSerLeu 232  
757 ATCGATGATCAACCGCTGCTTCGATGTCATGATCGCGCGGCAAGACAG 806  
233 ValAspGlyIleArgAlaThrAspValMetMetAlaGlyLysValAl 249  
807 TCTCGTATGCTGCTGCGGATGTCGGAAGGCTGCGGCTGCGCTCAATCC 856  
249 aValValCysGlyTyrGlyAspValGlyLysGlySerAlaAlaSerLeu 266  
857 GTGCCAAGGCTGCGGCTTATCATCACAGAACTCGACCCCAATCTCGCT 906  
266 rgGlyAlaGlyAlaArgValLysValThrGluValAspProIleCysAla 282  
907 CTCAGGCTGCTGGAAGGCTACAGGTCGCGCGCTGCGGCTGCGGAAAGT 956  
283 LeuGlnAlaAlaMetAspGlyPheGluValValLeuLeuGluAspVal 299  
957 CAAGATGCTCATATCTTCGTTACATGCACAGGAACTCCGATATCATCT 1006  
299 lAlaAspAla...IlePheIleThrThrGlyAsnLysAspValIle 315  
1007 CTGTTGATGATGCTGCGGCTGAGAGTGAAGGCTATTGTCGCGTAACATC 1056  
315 rgIleGluHisMetArgGluMetLysAspMetAlaIleValGlyAsnIle 331  
1057 GGCCACTTTCGATACGAAATTTGATACAGATGCGCTCATGAAATACCCAG 1106

```

|||||
332 GlyHisPheAspAsnGluIleGlnValAlaAlaLeuLysAsnHis..... 346
1107 CATCAAGCACATCCCAATCAAGCCAGAGAAATGACATGTTGGAAATCCCAAG 1156
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
347 ....LysTrpThrAsnIleLysAspGlnValAspMetIleGluMetProS 362
1157 ATGGCCACAGCTATCCCTCTCTGCTGAGGCCGCCCTTCTTAACCTTGGC 1206
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
362 erGlyAlaArgIleIleLeuLeuSerGluGlyArgLeuLeuAsnLeuGly 378
1207 TCGCGTACAGGTCACCCATCTTTCGTTATGTCATGTCATTCACAAACCA 1256
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 AsnAlaThrGlyHisProSerPheValMetSerAlaSerPheThrAsnG1 395
1257 GACACTCGCTCAGCTCGACCTCTACGAAAGAGAGAGAAATCTCGAAGA 1306
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
395 nValLeuAlaGlnIleGluLeuTrpThrLysGlyAlaGluTrpGlnProG 412
1307 AGGTTTACACTTCGAGACATCTCGATGAAGAAGTCGCGCTCCAC 1356
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
412 lyValTrpIleLeuProLysSerLeuAspGluLysValAlaArgLeuHis 428
1357 CTCGATCTCTCGATCTCCACCTTACAAAGCTTACACAGAGCAGCTGA 1406
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 LeuLysLysIleGlyValLysLeuThrThrLeuArgProAspGlnAlaG1 445
1407 CTACATCAACGTTCCAGTTGAGGGTCCCTTACAAGTCTGATGCTTACCGTT 1456
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
445 uTyrIleGlyValThrValGluGlyProPheLysSerAspHisTyrArgT 462
1457 AT 1458
||
462 Yr 462

```

seq\_name: pir2:AG3505

```

seq_documentation_block:
adenosylhomocysteinase (BC 3.3.1.1) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AG3505
R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova,
; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3505
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-481 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL53210.1; PID:gl7984086; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Map position: 1
C:Superfamily: adenosylhomocysteinase
C:Keywords: thioether hydrolase

```

```

alignment_scores:
Quality: 1343.00 Length: 475
Ratio: 3.562 Gaps: 4
Percent Similarity: 79.368 Percent Identity: 57.053

```

alignment\_block:

US-09-759-990-1 x AG3505 ..

Align seg 1/1 to: AG3505 from: 1 to: 481

```

34 GAGTACAGAAATTCGCACATCAACCTCCATGTTCTCGCGGTAAAGAACT 83
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
21 AspPheValValLysAspIleSerLeuAlaAspTrpGlyArgGluLe 37

```

```

84 TACCCTTGCTGAGAAGAAATGCCAGGTCTTATGTTCTTCTCGTGAGCGTT 133
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
37 uAspIleAlaGluThrGluMetProGlyLeuMetAlaAlaArgGluGlp 54
134 ATTCCGCTTCTTAAGCCATTAAGGGTGTGAGAATCTCTGGTTCCCTCCAC 183
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
54 heGlyLysSerGlnProLeuLysGlyAlaArgIleSerGlySerLeuHis 70
184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACACTCAGACTCTTGGTGC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
71 MetThrIleGlnThrAlaValLeuIleGluThrLeuLysValLeuGlyAl 87
234 TGATGTCAGATGGCTCTCTGCAACATCTCTCTACAAAGATACAGCGG 283
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
87 aGluValArgTrpAlaSerCysAsnIlePheSerThrGlnAspHisAla 104
284 CTGCTCTATCGTTGTCGGCCCAACAGGCACACAGCAGAAAGCCAGCGGT 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
104 laAlaAlaIleAlaAla.....ThrGlyThr..... 112
334 ATCCCAAGTCTTCGCTCGAAGGGGAAACACTCCAGAAATACATGGGAGAA 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
113 ...ProValPheAlaValLysGlyGluThrLeuGluGluTyrTrpThrTy 128
384 CACATACCCGCTCTCACATGGCCAGATGGTCAAGGCCACACAGCAGGTG 433
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
128 rThrAspGlnIlePheGlnTrpProAspGlyGluProSerAsnMetIleL 145
434 TCGATGATGGTGTGATGCTACACTCTCTCAAGGGGTTTCGAATTC 483
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
145 euAspAspGlyGlyAspAlaThrMetTyrIleLeuIleGlyAlaArgAla 161
484 GAAACAGCCGCTCTCTCCAGACCAAGAGCTGACACCTCGAATA 533
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
162 GluAlaGlyGluAspValLeuSerAsnProGlnSerGlu.....GluG1 176
534 CGCTGCGTCTTGTCTACACTCAAGCAGGTCTTCAACCAAGACAAAGCC 583
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 uGluValLeuPheAlaGlnIleLysLysArgMetAlaAlaThrProGlyP 193
584 ACTGGCACAGTTGCTCGCGGATGAACGGTGTTCGAGAGACAAACA 633
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 hePheThrLysGlnArgAlaAlaIleLysGlyValThrGluGluThrThr 209
634 ACAGTGTCCACCGCTCTACAGCTCGAAGAGGAGGCAAACTCCTCTT 683
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
210 ThrGlyValAsnArgLeuTyrGlnLeuGlnLysLysGlyLeuLeuProPh 226
684 CCCAGCCATCAACGTCACAGCAGCTGTTCACAAAGTCCAAAGTTCGATA 733
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
226 eProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspAsnL 243
734 TCTACGGCTCGCCCACTCCCTTATCGATGGTATCAACCGCTCTCCGAT 783
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
243 ystyrGlyCysLysGluSerLeuValAspGlyIleArgArgGlyThrAsp 259
784 GTCATGATCGCGGCAAGACAGCTCTCGTATGGTTCACGGCATGTCGG 833
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
260 ValMetMetAlaGlyLysValAlaValValCysGlyTyrGlyAspValG1 276
834 CAAGGGTCTCGTCAATCCCTCGGTGCCCAAGGCGCTCGCTTATCATCA 883
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
276 yLysGlySerAlaGlnSerLeuAlaGlyAlaGlyAlaArgValLysValT 293
884 CAGAACTCGACCCCAATCTGGCTCTCCAGGCTGCCATGGAAGGCTACCA 933
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
293 hrGluValAspProIleCysAlaLeuGlnAlaAlaMetAspGlyPheGlu 309
934 GTCCCGCGCATCGAGGAAGTCGTCAGAGTCTCGATATCTTCTCGTTACAT 983
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 ValValThrLeuAspAspAlaAlaSerThrAlaAspIleValValThrTh 326

```

984	CACAGGAACCTCGATATCATCTCTGTTGACATGATGCCACAGATGAAGG	1033
326	rThrGlyAsnLysAspValIleThrIleAspHisMetArgIysMetLysA	343
1034	ATAGGCTATTGTTCGGTACATCGGCCACTTCGATAACAGAAATTGATACA	1083
343	spMetCysIleValGlyAsnIleGlyHisPheAspAsnGluIleGlnVal	359
1084	GATGCCCTCATGAAATACCCAGGCATCAAGCACATCCCAATCAAGCCAGA	1133
360	AlaIleLeuArg.....AsnLeuLysTrpThrAsnValLysProGlu	373
1134	ATAGCACATGTGGGAATCCACAGATGCCACGCTATCCTCTCTTCTGTCG	1183
373	nValAspLeuIleGluPheProAspGlyLysArgLeuIleLeuLeuSerG	390
1184	AGGGCCGGCTTCTTAACCTTGGCTGGCTACAGGTACCCCATCTTTCGTT	1233
390	luGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSerPheVal	406
1234	ATGTCAATGTCATTCAACAACAGACACTCGCTCAGCTCGACCTCTACGA	1283
407	MetSerAlaSerPheThrAsnGlnValLeuGlyGlnIleGluLeuPheTh	423
1284	AAAGAGAGGAAATCTCGAGAAGAGGTTTACACACTTCCGAAGCATCTCG	1333
423	rArgThrAspAlaTyLysAsnGluValTyValLeuProLysHisLeuA	440
1334	ATGAGACAGTCTGCTCGCTCCACTCCGATCTCTCGATGTCACACCTTACA	1383
440	spGluLysValAlaArgLeuHisLeuAspLysLeuGlyAlaLysLeuThr	456
1384	AAGCTTACACAGAAGCGGTGACTACATCAACGTTCCAGTTGAGGCTCC	1433
457	ValLeuSerGluGluGlnAlaAlaTyIleGlyValThrProGlnGlyPr	473
1434	TTACAAGTCTGATGCTTACCGTTAT	1458
473	oPheLysSerGluHisTyArgTy	481

40 AGAATTGCCGACATCAACCTCCATGTTCTCGCGCGTGAAGAACTTACCCCT 89  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
7 LysValLysAspLysSerLeuAlaProPheGlyLysMetGlnL 23  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
90 TGCCTGAGAAGGAATCCAGGTCTTATGTTCTTCGFGAGCGTTATCCG 139  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
23 eSerGluAsnGluMetProGlyLeuMetArgLleArgGluGluTyrGlyL 40  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
140 CTTCTAAGCCATTGAAGGGTGCAGAAATCTCTGGTTCCTCCCTCCACATGACA 189  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
40 ysAspGlnProLeuLysAsnAlaLysIleThrGlyCysLeuHisMetThr 56  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
190 GTCCACACAGCCGCTCTCATCGACACACTCACGCTCTTGGTGCCTGATGT 239  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
57 ValGlyCysAlaLeuLeuIleGluThrLeuGlnLysLeuGlyAlaGlnI 73  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
240 CAGATGGGTCTTCGCAACATCTTCTACAGAAGATACAGCCGCTGCTG 289  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
73 eArgTyrCysSerCysAsnIleYrSerThrAlaAspTyrAlaAlaAla 90  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
290 CTATCGTTGTGGGCCCAACAGGCACACACAGAGAGCCAGCGGTATCCCA 339  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
90 laValLysAspGlnProLeuLysAsnAlaLysIleThrLeuGluLys 98  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
340 GTCTTCGCTCGAAGGCGGCAACACTCCCAAGATCTCGGAGAACACATA 389  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
99 ValPheAlaThrLysAsnGluThrLeuGluGluTyrTrpTyrCysValG 115  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
390 CCGCGGTCTCATAGCCAGCATGGT.....CAAGGCCACAGCAGGTG 433  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
115 uSerAlaLeuThrTrpGlyAspGlyAspAsnGlyProAspMetIleV 132  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
434 TCGATGATGGTGTGATGCTACACTCTCTCATCTCCAAGGGCTCGAA... 480  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
132 aLysAspGlyGlyAspAlaThrLeuLeuValHisLysGlyValGluTyr 148  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
481 .....TTCGAAACAGCGGTGCTGTTCCAGAGCCACAGAAAGTGA 521  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
149 GluLysLeuTyrGluGluLysAsnIleLeuProAspProGluLysAlaL 165  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
522 CAACCTCGAATCCGCTGCTTGTGTACACTCAACAGCGTCTTCAACC 571  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
165 sAsnGluGluGluArgCysPheLeuThrLeuLeuLysAsnSerIleLeu 182  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
572 AAGACAAAGAACCACTGGCACACAGTTGCTCGCGCATGAACGGTGTTC 621  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
182 ysAsnProLysLysTrpThrAsnIleAlaLysLysIleGlyValSer 198  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
622 GAAGACACACACAGGTGTCCACCGGCTCTACAGCTCGAGAGGAGGG 671  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
199 GluGluThrThrGlyValLeuArgLeuLysLysMetAspLysGlnAs 215  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
672 CAAACTCTCTTCCAGCCATCAACGTCACAGCGCTGTTCACAAAGTCCA 721  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
215 nGluLeuLeuPheThrAlaIleAsnValAsnAspAlaValThrLysGln 232  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
722 AGTTCGATAACATCTACGGTGGCGGCCTCCCTTATTCGATGGTATCAAC 771  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
232 ysrYrAspAsnValTyrGlyCysArgHisSerTyrThrAspGlyLeuMet 248  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
772 CGTGTCTCCGATGTCATGATCGGGCGGAAGACGCTCTCGTCATGGGTTA 821  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
249 ArgAlaThrAspPheLeuIleSerGlyLysIleValValIleCysGlyTy 265  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
822 CGCGCATGTCCGCAAGGGCTCGCTCAATCCCTCCGTGGCCAGCGCTC 871  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
265 rGlyAspValGlyLysGlyCysAlaSerSerMetLysGlyLeuGlyAla 282  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
872 GCGTTATCATCAGAACTCGACCCATCTCGCTCTCCAGGCTGCCATG 921  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
282 rgValTyrIleThrGluIleAspProIleCysAlaIleGlnAlaValMet 298  
: : : : : ||| ||| ||| ||| ||| ||| : : : : :  
922 GAAGGCTTACCAAGTTCGCCGATCGAGGAAGCTCGTCAAGGATGTGGAT 971

```

|||||.....|
299 GluGlyPheAsnValValThrLeuAspGluIleValAspLysGlyAspPh 315
972 CTTCCTTACATGCACGAAGAACTGCATATCATCTCTGTGTGACATGAG 1021
315 ePheIleThrCysThrGlyAsnValAspValIleLysLeuGluHisLeuL 332
1022 CCCAGATGACGATAAGGCTATTGTCGGTAACATCGGCCACTTCGATAAC 1071
332 euLysMetLysAsnAlaValValGlyAsnIleGlyHisPheAspAsp 348
1072 GAAATGTATACAGATGGCTCATGAAATACCCAGCATCAAGCACATCCC 1121
349 GluIleGlnValAsnGluLeuPheAsnTyrLysGlyIleHisIleGluAs 365
1122 AATCAACCCAGATPACGACATGTGGGAATCCAGATGCCACCGCTATCC 1171
365 nValLysProGlnValAspArgIleThrLeuProAsnGlyAsnLysIleI 382
1172 TCCTCTCTGCTGAGGCGCCTCTTACCTTGGCTGGCTACAGTCAAC 1221
382 leValLeuAlaArgGlyArgLeuLeuAsnLeuGlyCysAlaThrGlyHis 398
1222 CATCTCTTCTGTTATGTAATGTCATTACAAACACGACACTCGCTCAGCT 1271
399 ProAlaPheValMetSerPheCysAsnGlnThrPheAlaGlnLe 415
1272 CGACCTCTACGAAAGAGA.....GGAATCTCGAGAAGAGTTTACA 1315
415 uAspLeuTyrPginAsnLysAspThrAsnLysTyrGluAsnLysValTyrL 432
1316 CACTTCCGAGCATCTCGATGAAGAGTGCCTGCCCTCCACCTCGGATCT 1365
432 euLeuProLysHisLeuAspGluLysValAlaLeuTyrHisLeuLysLys 448
1366 CACGATGTCCTCATCAAGCTTACACAGAGGAGGCTGACTACATCAA 1415
449 LeuAsnAlaSerLeuThrGluLeuAspAsnGlnCysGlnPheLeuG1 465
1416 CTTCCAGTTGAGGTCTTACAAAGTCTGATGTACCGTAT 1458
465 yValAsnLysSerGlyProPheLysSerAsnGluTyrArgTyr 479
seq_name: pirl:A26583

```

```

seq_documentation_block:
adenosylhomocysteine (EC 3.3.1.1) - rat
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Nov-1988 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
R;Ogawa, H.; Gomi, T.; Mueckler, M.M.; Fujioka, M.; Backlund Jr., P.S.; Aksamit, R.R.; U
Proc. Natl. Acad. Sci. U.S.A. 84, 719-723, 1987
A;Title: Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from rat liver as de
A;Reference number: A26583; MUID:87118240
A;Accession: A26583
A;Molecule type: mRNA
A;Residues: 1-432 <OGA>
A;Cross-references: GB:M15185; NID:g202803; PIDN:AAAA0705.1; PID:g202804
R;Gomi, T.; Ogawa, H.; Fujioka, M.
J. Biol. Chem. 261, 13422-13425, 1986
A;Title: S-Adenosylhomocysteine from rat liver. Amino acid sequences of the peptides c
A;Reference number: A92560; MUID:87008564
A;Accession: A26191
A;Molecule type: protein
A;Residues: 76-94 <GOM>
A;Accession: B26191
A;Molecule type: protein
A;Residues: 104-121 <GO2>
A;Note: in a manner saturably inhibited by adenosine (a product) or adenine (a competitor
-79 and Cys-113, suggesting a role for these residues at the active site; however, neith
R;Merta, A.; Aksamit, R.R.; Kasir, J.; Cantoni, G.L.
Eur. J. Biochem. 229, 575-582, 1995

```

A;Title: The gene and pseudogenes of rat S-adenosyl-L-homocysteine hydrolase.  
A;Reference number: S69333; MUID:95262723  
A;Accession: S69333  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-432 <MER>  
A;Cross-references: EMBL:U14937; NID:g1223843; PIDN:AAA92043.1; PID:g1185363  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199  
C;Genetics:  
A;Introns: 10/1; 73/3; 99/1; 149/1; 186/3; 256/1; 285/2; 324/3; 389/3  
C;Function:  
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to ad  
C;Superfamily: adenosylhomocysteinease  
C;Keywords: homotetramer; NAD; thioether hydrolase  
F;214-243/Region: beta-alpha-beta NAD nucleotide-binding fold  
F;79,113/Active site: Cys #status predicted

## alignment\_scores:

Quality: 1289.00 Length: 490  
Ratio: 3.662 Gaps: 6  
Percent Similarity: 71.837 Percent Identity: 56.122

## alignment\_block:

US-09-759-990-1 x A26583

Align seg 1/1 to: A26583 from: 1 to: 432

```

1 ATGGCTTCAATCACTACTGCTGCTCCATTCGAGTACAGAAATTCGCGA 50
||||| ||| ||| |||.....|
1 MetAlaAspLysLeuPro.....TyrLysValAlaAs 11
51 CATCAACTCCATCTTCGCGCGTAAGAACTTACCTTGTGAGAAG 100
|||||.....| ||| |||.....|
11 PileGlyLeuAlaAlaThrGlyArgLysAlaLeuAspIleAlaGluAsn 28
101 AATGCCAGTCTTATGTTCTTCGTGAGCGTTATTCGCTTCTTAAGCCA 150
|||||.....| ||| |||.....|
28 LuMetProGlyLeuMetArgMetArgGluMetTyrSerAlaSerLysPro 44
151 TTGAAGGTCTCAGAACTCTGGTTCCTCCACATGACAGTCCAGACAGC 200
|||||.....| ||| |||.....|
45 LeuLysGlyAlaArgIleAlaGlyCysLeuHisMetThrValGluThrAl 61
201 CGTCCTCATCGACACTCACAGCTTGTGGTGTGATGTCAGATGGGCTT 250
|||||.....| ||| |||.....|
61 aValLeuIleGluThrLeuValAlaLeuGlyAlaGluValArgTyrPsrS 78
251 CTGCGACATCTTCTTACACAGATACAGCGCTGTGCTATCGTTGTC 300
|||||.....| ||| |||.....|
78 erCysAsnIlePheSerThrGlnAspHisAlaAlaAlaIleAla... 93
301 GGCCCAACAGGCACACAGAGAGCCAGCGGTATCCAGTCTTCGCGCTG 350
|||||.....| ||| |||.....|
94 .....LysAlaGlyIleProValPheAlaTr 102
351 GAAGGCGGAAACACTCCAGAAATACGGAGAACACATACCGCGCTCTCA 400
|||||.....| ||| |||.....|
102 pLysGlyGluThrAspGluGluTyrLeuTyrCysIleGluGlnThrLeuH 119
401 CATGCCAGATGTCAGGCCCCACAGCAGGTTCGTGATGATGGTGTAT 450
|||||.....| ||| |||.....|
119 isPheLysAspGlyPro...LeuAsnMetIleLeuAspAspGlyGlyAsp 134
451 GCTACACTCTCTCATCTCCAAAGGCTTCGAATTCGAAACAGCCGCTGT 500
|||||.....| ||| |||.....|
135 LeuThrAsnLeuIle..... 139
501 TCCAGAGCCACAGAGCTGACAACTCGAATACCGTCCGTTCTTCTGTA 550
|||||.....| ||| |||.....|
139 ..... 139
551 CACTCAAGCAGGTCTTCAACCAAGACAGAACCACTGGCACACA..... 594

```

426 ysProAspHisTyrArgTyr 432

seq\_name: pir2:T40763

seq\_documentation\_block:  
adenosylhomocysteine - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C;Accession: T40763  
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A  
submitted to the EMBL Data Library, March 1998  
A;Reference number: Z21948  
A;Accession: T40763  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-433 <LYN>  
A;Cross-references: EMBL:AL02072; PIDN:CAAL7833.1; GSPDB:GN00067; SPDB:SPBC8D2.18C  
A;Experimental source: strain 972h-; cosmid c8D2  
C;Genetics:  
A;Gene: SPDB:SPBC8D2.18C  
A;Map position: 2  
C;Superfamily: adenosylhomocysteine

alignment\_scores:  
Quality: 1274.00 Length: 475  
Ratio: 3.559 Gaps: 5  
Percent Similarity: 75.368 Percent Identity: 54.526

alignment\_block:

US-09-759-990-1 x T40763 ..

Align seg 1/1 to: T40763 from: 1 to: 433

37 TACAGAAATCCGACATCAACCTCCATGTTCTCGCGCGTAAAGAACTTAC 86  
|||||  
6 TyrLysValAlaAspIleSerLeuAlaAlaPheGlyArgLysGluLeuGl 22  
87 CCTGTCTGAGAAATGCCAGGTCTTATGTTCTTCTGAGGCTTAT 136  
|||||  
22 uileAlaGluAsnGluMetProGlyLeuileAlaValArgGluLysTyrA 39  
137 CCGTTCCTAAAGCCATTGAAGGTGTCTAGAAATCTCTGTTCTCCCTCCACATG 186  
||| |||  
39 lalysSerGlnProLeuLysGlyAlaAlaGlyLeuLysMet 55  
187 ACAGTCAGACAGCGCTCATCGAGACTCAGAGCTCTGTGCTGA 236  
56 ThrIleGlnThrAlaValLeuileGluThrLeuValAlaLeuGlyAlaGl 72  
237 TGTGAGATGGGCTTCTGCAACATCTTCTTACAAAGATACAGCGGTG 286  
||| |||  
72 uValThrTrpSerSerCysAsnIleTyrSerThrGlnAspHisAlaAla 89  
287 CTGCTATCTGTTCGGCCCAACAGGACACAGAGAGCCAGCGGTATC 336  
|||||  
89 laAlaileAlaAla .....ThrGlyVal 96  
337 CCAGTCTTCGCTGGAAGGGCGAAACACTCCAGAAATAC...TGGAGAA 383  
|||||  
97 ProValPheAlaTrpLysGlyGluThrGluGluLysLeuTrpCysIl 113  
384 CACATACCGGCTCTCACATGCGGATGTCAGAGCGCCACAGCAGCGGTG 433  
|||  
113 egluGlnGlnLeuLysSerPheProSerGlyLysProLeuAsnMetIleL 130  
434 TCGATCATGCTGTGATGCTACATCTCTCATCTCCAGGCGCTTCGAATTC 483  
|||  
130 euAspAspGlyLysPheLeuThrAlaLeuValHisGluArgHis..... 144  
484 GAAACAGCGGTGCTCTCCAGAGCCACAGAGAGCTGACAACTCGAATA 533  
|||||  
145 .....ProGlu..... 146

140 .....HisthrLysHis 143  
595 .....GTTGCTCGCGGATGACGGTTCGGAAGAGACAAACAGG 638  
|||  
144 ProGlnLeuLeuSerGlyIleArgGlyLeuSerGluGluThrThrGl 160  
639 TGTCCACCGCTCTACAGCTCCAGAGGAGGCGCAAACTCTCTTCCAG 688  
|||  
160 yValHisAsnLeuTyrLysMetMetAlaAsnGlyLeuLysValProA 177  
689 CCATCAACGTCAACAGCGCTGTACAAAGTCCAAAGTTCGATACATCTAC 738  
|||||  
177 laileAsnValAsnAspSerValThrLysSerLysPheAspAsnLeuTyr 193  
739 GGTGCGCGCACCTCTTATCGATGATGATCAACCGTCTCTCCGATGTCAT 788  
|||||  
194 GlycArgGluSerLeuileAspGlyLeuLysArgAlaThrAspValme 210  
789 GATCGCGGCAACAGACTCTGTCATGGTTAGCGGATGTCGCGAAGG 838  
|||||  
210 tileAlaGlyLysValAlaValAlaGlyTyrGlyAspValGlyLysG 227  
839 GCTCGCTCAATCCCTCGGCGCAAGCGCTCGCTTATCATCACAGAA 888  
|||||  
227 lYcysAlaGlnAlaLeuArgGlyPheGlyAlaArgValIleleThrGlu 243  
889 CTCGACCAATCTCGCTCTCCAGGTCGCCATGAGGCTACCGGTCG 938  
|||||  
244 ileAspProIleAsnAlaLeuGlnAlaAlaMetGluGlyTyrGluValth 260  
939 CCGATCCGAGAACTCGTCAAGGATGTCGATATCTTCTATCATCACAG 988  
|||||  
260 rThrMetAspGluAlaCysLysGluGlyAsnIlePheValThrThrG 277  
989 GAAACTGGGATATCATCTCTGTGATGATGATGCGCCGATGAGGATAG 1038  
|||  
277 lYcysValAspIleleleLeuGlyArgHisPheGluGlnMetLysAspSp 293  
1039 GCTATGTCGGTACATCGGCACCTCGATACAGAAATGATACAGATGG 1088  
294 AlaIleValCysAsnIleGlyHisPheAspValGluileAspValLysTr 310  
1089 COTCATGAAATCAACCCAGGATCAAGCACATCCCAATCAAGCCAGAAATAG 1138  
|||  
310 lueAsnGlu...AsnAlaValGluLysValAsnIleLysProGlnValA 326  
1139 ACATGTCGGAATTCAGATGCCAGCTATCTCTCTTCTGCTGAGGC 1188  
|||  
326 spArgTyrLeuLeuLysAsnGlyHisArgIleleleLeuAlaGluGly 342  
1189 CGCTCTTAACTTGGCTGCGCTACAGGTCACCCATCTTCTGTTATGTC 1238  
|||||  
343 ArgLeuValAsnLeuGlyCysAlaMetGlyHisProSerPheValMetSe 359  
1239 AATGTCATTCAAAACAGACACTCCCTCGCTAGCTCGAGCTTACGAAAGA 1288  
|||  
359 rAsnSerPheThrAsnGlnValMetAlaGlnIleGluLeuThrPheHisP 376  
1289 GAGGAATCTCGAAGAAGGTTTACACATCTCCGAGCATCTCCGAGCATCTCGATGAA 1338  
|||  
376 roAspLysTyrProValGlyValHisPheLeuProLysLysLeuAspGlu 392  
1339 GAATGCTGCTCGCTCCACCTCGGATCTCTCGATGTCACCTTACAAAGCT 1388  
|||||  
393 AlaValAlaGluAlaHisLeuGlyLysLeuAsnValLysLeuThrLysLe 409  
1389 TACACAGAACGAGCTCACTACATCAAGTTCAGGTTGAGGCTCCTTACA 1438  
|||||  
409 uThrGluLysGlnAlaGlnTyrLeuGlyMetProIleAsnGlyProPheL 426  
1439 AGTCTGATGCTTACCGTTAT 1458  
|||



534 CCCTCGCTTCTTCTGCTACACTCAAGCAGGTCTTCAACCAAGACAAGAAC 583  
 146 ..... 146  
 584 ACTGGCACACAGTTGCTGCGGCGATGAACGGTGTTCGGAAGACAACA 633  
 147 ..... 147  
 634 ACAGTGTCCACGCCCTCTACCACTCGAGAGGAGGCAAACTCCTCTT 683  
 160 ThrGlyValHisAsnLeuTyrLysMetPheLysGluAsnLysLeuLysVa 176  
 684 CCAGCCATCAACCTCAACGACGCTGTTCACAAAGTCCAAGTTCGATAACA 733  
 176 IProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspAsnL 193  
 734 TCTACGGCTCGCGCCACTCCCTTATCGATGCTATCAACCGTCTCCGAT 783  
 193 euPheGlyCysLysGluSerLeuValAspGlyIleLysArgAlaThrAsp 209  
 784 GTCATGATCGCGCGGCAAGACGCTCGCTCATGGGTACGCGCATGTCGG 833  
 210 ValMetIleAlaGlyLysValAlaValAlaGlyPheGlyAspValGI 226  
 834 CAAGGCTCGCTCAATCCCTCGGTGGCCAAAGCGCTCGGTTCATCA 883  
 226 yLysGlyCysSerThrSerLeuArgSerGlnGlyAlaArgValIleValT 243  
 884 CAGAACTCGACCAATCTCGCTCTCCAGCTGCCATGGAAGGCTACCAG 933  
 243 hrGluValAspProIleAsnAlaLeuGlnAlaAlaMetaspGlyPheGlu 259  
 934 GTCGCCCGCATCGAGAGTCTCAAGGATGCGATATCTTCGTTACATG 983  
 260 ValThrThrMetGluGluAlaValLysGluGlyGlnIlePheValThr 276  
 984 CACAGGAACCTGCGATATCATCTCTGTCACATGATGCCCAAGG 1033  
 276 rThrGlyCysArgAspIleIleArgGlyGluHisPheAsnGluMetLysG 293  
 1034 ATAAGGCTATTGTCGGTAACTCGGCCACTTCGATAACGAAATTCATACA 1083  
 293 luAspSerIleValCysAsnIleGlyHisPheAspValGluIleAspVal 309  
 1084 GATGGCTCATGAAATACCCAGGCATCAAGCACATCCCAATCAAGCCAGA 1133  
 310 ...AlaTrpLeuLysAlaAsnAlaLysAspValValAsnIleLysProGI 325  
 1134 ATAGCAGATGTCGGAATTCAGATGCGCCAGCTATCCCTCTTCGCTG 1183  
 325 nValAspArgTyrGluLeuLysAsnGlyArgHisIleIleLeuLeuAla 342  
 1184 AGGCGCGCTTCTTAACCTTGGCTGCGGTACAGGTCAAGTTCCTTTCGTT 1233  
 342 spGlyArgLeuValAsnLeuGlyCysAlaThrGlyHisProSerPheVal 358  
 1234 ATGTCATGTCATTCACAAACAGACACTCGCTCAGCTCGACCTCTACGA 1283  
 359 MetSerCysSerPheThrAsnGlnValLeuAlaGlnIleAlaLeuTrpTh 375  
 1284 AAAGAGAGGAATCTCGAGAGAAGGTTTACACACTTCGAGCATCTCG 1333  
 375 rAspAsnThrSerTyrProLeuGlyValHisMetLeuProLysLysLeuA 392  
 1334 ATGAAGAAGTCGCTCGCTCCACCTCGATCTCTCGATGTCACCTTACA 1383  
 392 spLugluValAlaArgLeuHisLeuGlyLysLeuGlyValLysLeuThr 408  
 1384 AGCTTACAGAGGAGGCTGACTACATCAACCTTCAGTTGAGGTCC 1433  
 409 ThrLeuThrSerValGlnSerAspTyrLeuGlyIleProValAspGlyPr 425

1434 TTACAAGTCTGATGCTTACCGTTAT 1458  
 425 OTYrLysAlaAspHisTyrArgTyr 433

us-09-759-990-1.n2p.rpr

Mon Aug 19 15:16:21 2002



alignment\_scores:  
 Quality: 1550.50 Length: 477  
 Ratio: 3.945 Gaps: 3  
 Percent Similarity: 82.390 Percent Identity: 62.264

alignment\_block:  
 US-09-759-990-1 x AAY87873 ..

Align seg 1/1 to: AAY87873 from: 1 to: 495

```

34 GAGTACAGAAATGCCGACATCAACCTCATCTTCGCGCGGTAAGGAAC 83
20 AspPheLysIleAlaAspLeuSerLeuAlaAspPheGlyArgLysGluLe 36
84 TACCCCTGCTGAGAGAAATGCCAGGCTCTTATGTTCTGTCGAGCGTT 133
1 uAlaGluValGlnProLeuLysGlyAlaArgIleSerGlySerLeuHis 69
36 uArgIleAlaGluHisGluMetProGlyLeuMetSerLeuArgArgGlu 53
134 ATTCGGCTTCTAAGCCATTGAAGGGTCTCAGAACTCTCTGTTCCCTCC 183
53 yAlaGluValGlnProLeuLysGlyAlaArgIleSerGlySerLeuHis 69
184 ATTCAGAGTCCAGACGCGCTTCATCGAGACATCAAGCTCTTGGTGC 233
70 MetThrValGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 86
234 TGATGTCAGATGGCTTCGCGCAACATCTTCTACACAGATACAGCGG 283
86 aGluValArgTrpAlaSerCysAsnIlePheSerThrGlnAspHisAla 103
284 CTGCTGCTATCTGTCGCGCCACAGGACACACAGAGACCCAGCCGGT 333
103 laAlaAlaValValGlyProHisGlyThrProAspGluProLysGly 119
334 ATCCAGTCTTCGCTCGGAGGCGCAACACTCCAGAACTACTGGGAAA 383
120 ValProValPheAlaTrpLysGlyGluThrLeuGluGluTrpTrpAl 136
384 CACATACCCGCTCTCAGATGCGCAGATGTCTCAAGGCCCA...CAGCAG 430
136 aAlaGluGlnMetLeuThrTrpProAspProAspLysProAlaAsnMetI 153
431 TTGTCGATGATGCTGTGATGCTACACTCTCTCATCTCCAGGGCTCGAA 480
153 leuAspAspGlyGlyAlaGluSerValLysGlyValLeuArgGlyMetGln 169
481 TTCGAACACGCGCTGCTGTCAGAGCCCAACAGAGCTGACACCTCGA 530
170 TyrGluLysAlaGlyValProAlaGluGluAspAspProAlaGlu 186
531 ATACCGCTCGGTTCTGTCATCACTCAAGCAGGCTTCAACCAACAGAGA 580
186 uTrpLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLys 203
581 ACCACTGCGCACAGTGTGCTCGCGGATGAACGCTGTTCCGAGAGAGACA 630
203 sPlyTrpThrLysIleAlaGluSerValLysGlyValThrGluGluThr 219
631 ACAACAGGTTCCACCGCTCTACAGCTCGAGAGGAGGCGCAACCTCT 680
220 ThrThrGlyValLeuArgLeuTyrGlnPheAlaAlaGlyAspLeuAl 236
681 CTTCGAGCATCAACGTCACACGCTGTGTACAAAGTCCAGTTCGATA 730
236 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAsp 253
731 ACATCTACGCTCGCGCCACTCTCTATCGATGATGATCAACCGCTTCC 780
253 snLysTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 269
781 GATGTCATGATCGCGCGCACAGACAGCTCTCTGTCATGGGTTACGGCGAT 830
270 AspAlaLeuIleGlyGlyLysValLeuIleCysGlyTyrGlyAspVa 286
  
```

```

831 CGCAAGGGTGCCTCAATCCCTCCGTGGCCAGCGCTCGGTTATCA 880
286 lGlyLysGlyCysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 303
881 TCACAGAACTCGACCCCAATCTGCGCTCTCCAGGCTGCCATGGAAGCTAC 930
303 alThrGluIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 319
931 CAGTCCGCGCATCGAGGAAGTCTCAAGGATGTCGATATCTTCTGTTAC 980
320 AspValValThrValGluGluAlaIleGlyAspAlaAspIleValValTh 336
981 ATGCACAGAACTCGCATATCATCTCTGTTGATGATGATGATGATGATGAT 1030
336 rAlaThrGlyAsnLysAspIleIleMetLeuGluHisIleLysAlaMetL 353
1031 AGGATAGGCTATTGTCGTAACATCGGCACTCGGCTTCTGATAACGAAATTGAT 1080
353 yAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 369
1081 ACAGATGGCTCATGAATACCCAGGATCAAGCACATCCCCAATCAAGCC 1130
370 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysPr 385
1131 AGAATAGGATGTTGGAATTCCTCAGAT...GGCCACGCTATCTCTCTTC 1177
385 oGlnValAspLeuTrpPheGlyAspThrGlyArgSerIleIleValL 402
1178 TTGCTGAGGCGCGCTTCTTAACCTTGGCTGCTGCTGCTGCTGCTGCTGCT 1227
402 euserGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 418
1228 TTGCTTATGTCAATGTCTTCAACCAACAGACACTCGCTCAGCTCGACCT 1277
419 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 435
1278 CTAGAAAGAGAGAAATCTCGAGAGAGAGGTTTACACACTTCCGAGAC 1327
435 uTrpThrLysAsnAspGluTyrAspAsnGluValTyrArgLeuProLysH 452
1328 ATCTCGATGAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377
452 isLeuAspGluLysValAlaArgIleHisValGluAlaLeuGlyHis 468
1378 CTTCACAAAGCTTACACAGAGAGGCTGACTACATCAACGTTTCCAGTTGA 1427
469 LeuThrLysLeuThrLysGluGlnAlaGluTyrLeuGlyValAspValG 485
1428 GGTCTCTTACAAGTCTGATGTTACCGTTAT 1458
485 uGlyProTyrLysProAspHisTyrArgTyr 495
  
```

seq\_name: /SDSS/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAG40086

seq\_documentation\_block:  
 ID AAG40086 standard; Protein; 485 AA.  
 XX AAG40086;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 49691.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.

XX	25-FEB-2000; 2000EP-0301439.	
PF		
XX	25-FEB-1999; 99US-0121825.	99US-0142803.
PR	05-MAR-1999; 99US-0123180.	99US-0142920.
PR	09-MAR-1999; 99US-0123548.	99US-0142977.
PR	23-MAR-1999; 99US-0125788.	99US-0143542.
PR	25-MAR-1999; 99US-0126264.	99US-0143624.
PR	29-MAR-1999; 99US-0126785.	99US-0144005.
PR	01-APR-1999; 99US-0127462.	99US-0144085.
PR	08-APR-1999; 99US-0128234.	99US-0144086.
PR	16-APR-1999; 99US-0128714.	99US-0144325.
PR	19-APR-1999; 99US-0129845.	99US-0144331.
PR	21-APR-1999; 99US-0130077.	99US-0144332.
PR	23-APR-1999; 99US-0130449.	99US-0144333.
PR	28-APR-1999; 99US-0130891.	99US-0144334.
PR	30-APR-1999; 99US-0131449.	99US-0144335.
PR	30-APR-1999; 99US-0132048.	99US-0144352.
PR	04-MAY-1999; 99US-0132407.	99US-0144352.
PR	05-MAY-1999; 99US-0132484.	99US-0144884.
PR	06-MAY-1999; 99US-0132485.	99US-0144814.
PR	06-MAY-1999; 99US-0132486.	99US-0145086.
PR	07-MAY-1999; 99US-0132487.	99US-0145088.
PR	11-MAY-1999; 99US-0132863.	99US-0145085.
PR	14-MAY-1999; 99US-0134256.	99US-0145087.
PR	14-MAY-1999; 99US-0134218.	99US-0145089.
PR	14-MAY-1999; 99US-0134219.	99US-0145192.
PR	14-MAY-1999; 99US-0134221.	99US-0145145.
PR	18-MAY-1999; 99US-0134370.	99US-0145218.
PR	19-MAY-1999; 99US-0134768.	99US-0145224.
PR	20-MAY-1999; 99US-0134941.	99US-0145276.
PR	21-MAY-1999; 99US-0135124.	99US-0145913.
PR	21-MAY-1999; 99US-0135353.	99US-0145918.
PR	24-MAY-1999; 99US-0135629.	99US-0145919.
PR	25-MAY-1999; 99US-0136021.	99US-0145951.
PR	27-MAY-1999; 99US-0136392.	99US-0146386.
PR	28-MAY-1999; 99US-0136782.	99US-0146389.
PR	01-JUN-1999; 99US-0137222.	99US-0147038.
PR	03-JUN-1999; 99US-0137528.	99US-0147204.
PR	04-JUN-1999; 99US-0137502.	99US-0147302.
PR	07-JUN-1999; 99US-0137724.	99US-0147192.
PR	08-JUN-1999; 99US-0138094.	99US-0147260.
PR	10-JUN-1999; 99US-0138540.	99US-0147303.
PR	14-JUN-1999; 99US-0139119.	99US-0147416.
PR	16-JUN-1999; 99US-0139452.	99US-0147493.
PR	17-JUN-1999; 99US-0139453.	99US-0147935.
PR	18-JUN-1999; 99US-0139492.	99US-0148171.
PR	18-JUN-1999; 99US-0139454.	99US-0148319.
PR	18-JUN-1999; 99US-0139455.	99US-0148341.
PR	18-JUN-1999; 99US-0139456.	99US-0148565.
PR	18-JUN-1999; 99US-0139457.	99US-0148684.
PR	18-JUN-1999; 99US-0139458.	99US-0149368.
PR	18-JUN-1999; 99US-0139459.	99US-0149175.
PR	18-JUN-1999; 99US-0139460.	99US-0149426.
PR	18-JUN-1999; 99US-0139461.	99US-0149722.
PR	18-JUN-1999; 99US-0139462.	99US-0149723.
PR	18-JUN-1999; 99US-0139463.	99US-0149929.
PR	18-JUN-1999; 99US-0139464.	99US-0149902.
PR	18-JUN-1999; 99US-0139750.	99US-0149930.
PR	21-JUN-1999; 99US-0139817.	99US-0150566.
PR	22-JUN-1999; 99US-0139859.	99US-0150884.
PR	23-JUN-1999; 99US-0140353.	99US-0151065.
PR	24-JUN-1999; 99US-0140354.	99US-0151066.
PR	24-JUN-1999; 99US-0140695.	99US-0151080.
PR	29-JUN-1999; 99US-0140823.	99US-0151303.
PR	30-JUN-1999; 99US-0140991.	99US-0151303.
PR	01-JUL-1999; 99US-0141287.	99US-0151930.
PR	01-JUL-1999; 99US-0141842.	99US-0152363.
PR	02-JUL-1999; 99US-0142154.	99US-0153070.
PR	06-JUL-1999; 99US-0142055.	99US-0153758.
PR		99US-0154018.
PR		99US-0154039.
PR		99US-0154779.
PR		99US-0155139.
PR		99US-0155486.
PR		
PR	08-JUL-1999;	
PR	09-JUL-1999;	
PR	12-JUL-1999;	
PR	13-JUL-1999;	
PR	14-JUL-1999;	
PR	15-JUL-1999;	
PR	16-JUL-1999;	
PR	16-JUL-1999;	
PR	19-JUL-1999;	
PR	19-JUL-1999;	
PR	19-JUL-1999;	
PR	19-JUL-1999;	
PR	20-JUL-1999;	
PR	20-JUL-1999;	
PR	20-JUL-1999;	
PR	21-JUL-1999;	
PR	21-JUL-1999;	
PR	21-JUL-1999;	
PR	22-JUL-1999;	
PR	22-JUL-1999;	
PR	22-JUL-1999;	
PR	23-JUL-1999;	
PR	23-JUL-1999;	
PR	26-JUL-1999;	
PR	27-JUL-1999;	
PR	27-JUL-1999;	
PR	27-JUL-1999;	
PR	02-AUG-1999;	
PR	02-AUG-1999;	
PR	03-AUG-1999;	
PR	04-AUG-1999;	
PR	04-AUG-1999;	
PR	05-AUG-1999;	
PR	05-AUG-1999;	
PR	06-AUG-1999;	
PR	06-AUG-1999;	
PR	09-AUG-1999;	
PR	09-AUG-1999;	
PR	10-AUG-1999;	
PR	11-AUG-1999;	
PR	12-AUG-1999;	
PR	13-AUG-1999;	
PR	13-AUG-1999;	
PR	16-AUG-1999;	
PR	17-AUG-1999;	
PR	18-AUG-1999;	
PR	20-AUG-1999;	
PR	20-AUG-1999;	
PR	23-AUG-1999;	
PR	23-AUG-1999;	
PR	25-AUG-1999;	
PR	26-AUG-1999;	
PR	27-AUG-1999;	
PR	27-AUG-1999;	
PR	30-AUG-1999;	
PR	31-AUG-1999;	
PR	01-SEP-1999;	
PR	07-SEP-1999;	
PR	10-SEP-1999;	
PR	13-SEP-1999;	
PR	15-SEP-1999;	
PR	16-SEP-1999;	
PR	20-SEP-1999;	
PR	22-SEP-1999;	
PR	23-SEP-1999;	

Mon Aug 19 15:16:18 2002

```

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
Quality: 1537.00      Length: 487
Ratio: 3.901          Gaps: 4
Percent Similarity: 80.903 Percent Identity: 62.012

alignment_block:
US-09-759-990-1 x AAG40086 ..

Align seg 1/1 to: AAG40086 from: 1 to: 485

19 ACTGGTGCTCCATTCGAGTACAGNAATTCGCGACATCAACCTCCATGTTCT 68
8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24
69 CGCGCGTAAGGAAGTACCTTACCTGTCGAGAGGAAATCCAGGCTTATGG 118
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMetA 41
119 TTCCTGCTGAGCGTTATCCGCTTCTAAGCCATTGAAGGTTGTCAGATC 168
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile 57
169 TCTGGTTCCTCCACATGACAGTCCAGACGCGCTCTCATCGAGACACT 218
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
219 CACAGCTCTTGGTGCTGATGTCAGATGGGCTTCCTCGCAACATCTCTCTA 268
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerT 91
269 CACAAGATACAGCGCTCTGCTATCTGTTGTCGGCCCAACAGGCACACCA 318
91 hrGlnAspHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 105
319 GAGAGCCAGCGGTATCCAGTCTTCGCTTGGAGGCGGAAACACTCCC 368

```

```

106 .....ValPheAlaTriPlysGlyGluThrLeuG1 115
369 AGAATACTGGGAGACACATACCGCGCTCTCACATGGCCAGATGGTCAAG 418
115 nGluTyrTriPTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 132
419 GCCACAGCAGAGGTTGTCGATGATGGTGGTGTATCATCACTCCTCATCC 468
132 lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148
469 AAGGGCTTCGAA.....TTCGAAACAGCGGTGCTGTTCCAGA 506
149 GluGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProAs 165
507 GCCAACAGAGAGCTGACAACTCGAATACCGCTGCGTTCCTGCTACATCA 556
165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLeL 182
557 AGCAGGTCTTCAACCAAGACAAAGAACCTGGCACACAGTTGCTGCCGCG 606
182 ysGluGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 198
607 ATGACAGGTGTTCCGAGAGACAAACAGGTGTCCACCGCTCTACCA 656
199 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyrG1 215
657 GCTCGAGAAGGAGGCAAACTCTCTTCCAGCCATCAACGTCACACGACG 706
215 nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsps 232
707 CTGTTACAAAGTCCAAAGTTCGATACATCTACGGCTCGCGCTCCCTCTT 756
232 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
757 ATCGATGTATCAACCGCTGCTTCGATGTATGATCGCGCGGCAAGACAC 806
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
807 TCTCGTCATGGTTACGGGATGTCGGCAAGGCTGCGCTCAATCCCTCC 856
265 aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaAlaMetL 282
857 GTGGCCAAAGCGCTCGGTTATCATCACAGACTCGACCCCAATCTCGCT 906
282 ysThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
907 CTCAGGCTGCCATGGAAGCTACCGAGTCCCGCGCATCGAGGAAGTCGT 956
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 315
957 CAAGGATGTCGATATCTTCGTTACATGCACAGAAAGCTGCGGATATCATCT 1006
315 lSerGluAlaAspIlePheValThrThrThrGlyAsnLysAspIleIleM 332
1007 CTGTTACATGATGGCCCGCATGAAGTAAAGCTATTGTCGGTAACATC 1056
332 etValAspHisMetArgLysMetLysMetLysAsnAlaIleValCysAsnIle 348
1057 GGCACCTTCGATAACGAAATTTGATACAGATGGCTCATGAAATACCCAGG 1106
349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrProGl 365
1107 CATCAAGCACATCCCAATCAAGCCAGATACGACATGTCGGAATTCACG 1156
365 yValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProG 382
1157 ATGCCACGCT...ATCCTCTCTTCTGTCGAGGCGCGCTCTTTAACCTT 1203
382 iuThrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
1204 GGCTGCGCTACAGTCAACCATCTTCTGTTATGTCATGTATGTCATCACAA 1253

```







```

707 CTGTTACAAAGTCCAAGTTCCGATAACATCTACGGCTGCGGCGCACTCCCTT 756
255 :rValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 271
757 ATCGATGGTATCAACCGTGTCTCCGATGTATGCGGCGCAAGACAGC 806
272 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 288
807 TCTGTCATGGTTAGCGGATGTGCGCAAGGGGTGCGCTCAATCCCTCC 856
288 aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetL 305
857 GTGGCCAAAGCGCTCGCTTATCATCACAGAACTCGACCCCAATCTGCCT 906
305 ysThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 321
907 CTCAGGCTGCCATGGAAGCTACCAAGTCCGCGCATCGAGGAAGTCGT 956
322 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 338
957 CAAGGATGTCGATATCTGTTACATGCACAGAGAACTGGGATATCATCT 1006
338 lserGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 355
1007 CTGTTGACATGATGCGCCAGATGAAGGATATTGTGCGTAACATC 1056
355 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 371
1057 GGCACCTCGATAACGAATGATACAGATGGCTCATGAATACCCAGG 1106
372 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrProG 388
1107 CATCAAGCACATCCCAATCAAGCAATACGACATGTGGGAATTCCTCAG 1156
388 yValLysArgIleThrIleLysProGlnThrAspArgTyrPheProG 405
1157 ATGGCCACGCT...ATCCTCTCTTCTGTGAGGCGCGCTTCTTAACCTT 1203
405 luthrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 421
1204 GGCCTCGCTACAGGTCACCCATCTTTCGTTATGTCAATGTCATTCACAA 1253
422 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 438
1254 CCAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAAATC 1297
438 nGlnValIleAlaGlnLeuGluLeuThrPasnGluLysAlaSerGlyLysT 455
1298 TCGAGAGAAGGTTTACACACTTCGGAAGCATCTCGATGAAGAAGTCGCT 1347
455 yrGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 471
1348 CGCCTCCACCTCGGATCTCGATGTCACCTTACAAAGCTTACACAGAA 1397
472 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAs 488
1398 CGAGGCTGACTACATCAAGTTCACAGTTCAGGTCCTTCAAGTCTGATG 1447
488 pGlnSerAspTyrValSerIleProIleGluGlyProTyrLysProProH 505
1448 CTRACCGTTAT 1458
505 lstyArgTyr 508

```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:AA26500

seq\_documentation\_block:

ID AAR26500 standard; Protein: 486 AA.

XX AC AAR26500;

XX DT 10-MAR-1993 (first entry)

XX

Prod. of Nicotiana tabacum gene expressing at floral differentiation.  
Flower; induction.

Nicotiana tabacum.

JP04258292-A.

14-SEP-1992.

14-FEB-1991; 91JP-0020702.

14-FEB-1991; 91JP-0020702.

(NISB) JAPAN TOBACCO INC.

WPI: 1992-354683/43.

N-PSDB; AAQ29419.

Gene expressing at floral differentiation for flowering control -  
obtd. from culture of floral axis epithelium cells of Nicotiana  
tabacum, for introduction into other plants

Disclosure; Page 6; 8pp; Japanese.

The protein sequence was deduced from the DNA sequence of a clone  
obtd. by screening a cultured epithelial axis cDNA library with probes  
from cultured and untreated mRNA. The clone obtd. expressed its  
protein only at floral differentiation. The gene can be introduced  
into other plants or can be suppressed by an antisense technique for  
the control of flowering of plants.

Sequence 486 AA;

alignment\_scores:

Quality: 1526.00 Length: 482

Ratio: 3.953 Gaps: 4

Percent Similarity: 80.083 Percent Identity: 63.071

alignment\_block:

US-09-759-990-1 x AAR26500 ..

Align seg 1/1 to: AAR26500 from: 1 to: 486

```

34 GAGTACAGAAATGCCGACATCAACCTCCATGTTCTCGGCGCTTAAGGAAC 83
14 GluTyrLysValLysAspMetSerGlnAlaAspPheGlyArgLeuGluI 30
84 TACCCTTGTGAGAAGAAATGCCAGTCTTATGGTTCTTCGTGAGCGTT 133
30 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgThrGluP 47
134 ATTCGCTTCTTAGCCATTCAGGGTGTCAAGATCTCTGGTTCCTCCAC 183
47 heGlyProSerGlnProPheLysGlyAlaLysIleThrGlySerLeuHis 63
184 ATGACAGTCCAGACAGCGCTCTCATCGACACTCACAGCTCTTGGTGC 233
64 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 80
234 TGATGTCAGATGGCTTCCTGCAACATCTTCTTACACAGATACAGCG 283
80 aGluValArgTyrCysSerCysAsnIlePheSerThrGlnAspHisAlaA 97
284 CTGCTGCTATCTGTTGCGGCCCAACAGGCACACCAGAGAGCCAGCGGT 333
97 laAlaAlaIleAlaArgAspSerAlaAla..... 106
334 ATCCAGCTTTCGCTGGAGGCGGAAACACTCCCGAATACTCGGAGAA 383
107 .....ValPheAlaTyrLysGlyGluThrLeuGlnGluTyrTrpCp 121

```

384 CACATACCGCGCTCTCACATGCGCCAGTGGTCAAGCCGACAGCAGGTTG 433  
121 sThrGluArgAlaLeuAspTrpGlyProGlyGlyProAspLeuIleV 138  
434 TCATGATGGTGGTGGTACATCCCTCATCTCCAAAGGCGTCTC..... 477  
138 alaAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 154  
478 .....GAATTCGAACAGCGCGTGTTCAGAGCCCAACAGAGCTGA 521  
155 GluGluGluPheAlaLysAsnGlyThrIleProAspProAsnSerThrAs 171  
522 CAACCTCGAATACCGCTGGTTCCTGCTACACTCAAGCAGGCTTCAACC 571  
171 pAsnAlaGluPheGlnLeuValLeuThrIleLeuLysGluSerLeuLysT 188  
572 AAGACAAGAACCCATGGCACAGTGTGTCGCGGCATGAACGGTGTTC 621  
188 hrAspProLeuLysThrLysMetLysGluArgLeuValGlyValSer 204  
622 GAAGAGACAGAACAGGTGTCCACCGCTCTACAGCTCGAAGAGGAGG 671  
205 GluGluThrThrGlyValLysArgLeuThrIleMetGlnAlaAsnG 221  
672 CAACCTCTCTCTCCAGCCATCAACGCTCAACGCTGTTCACAAAGTCCA 721  
221 yThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSerL 238  
722 AGTTCGATACATCTACGCTGCGCGCATCTCCCTTATCGATGTATCAAC 771  
238 ysPheAspAsnLeuTyGlyCysArgHisSerLeuProAspGlyLeuMet 254  
772 CGTGTCTCGATGTCATGTCGCGGCAGACAGCTCTCGTCATGGTTA 821  
255 ArgAlaThrAspValMetIleAlaGlyLysValAlaLeuValAlaGlyT 271  
822 CGCGCATGTCGCGCAAGGCTCGCTCAATCCCTCGTGGCGCAAGGCGCTC 871  
271 rGlyAspValGlyLysGlyCysAlaAlaLeuLysGlnAlaGlyAlaA 288  
872 GGTATATCATCAGAACATCGACCCATCTCGCTCTCCAGGCTGCCATG 921  
288 rgValIleValThrGluIleAspProIleCysAlaLeuGlnAlaThrMet 304  
922 GAAGGCTACAGCTCGCGCATCTCGAGGAGTGTCTCAAGGATGTCTGATAT 971  
305 GluGlyLeuGlnValLeuThrLeuGluAspValValSerAspValAspI 321  
972 CTTCTGTTACATGCACAGGAAGTGGCATATCATCTCTGTTGCATGATGG 1021  
321 ePheValThrThrGlyAsnLysAspIleIleMetValAspHisMetA 338  
1022 CCAGATGAAGGATAAGGCTATTGTGCGTAACTACATCGCCACTTCGATAAC 1071  
338 rgLysMetLysAsnAsnAlaIleValCysAsnIleGlyHisPheAspAsn 354  
1072 GAATTCGATACAGATGGCTCTCATGAATCCAGGATCAACAGCACATCCC 1121  
355 GluIleAspMetLeuGlyLeuGluThrTyProGlyValLysLysIleTh 371  
1122 AATCAAGCCAGATACCATGTCGGAATTCCTCCAGATGGCCACGCT...A 1168  
371 rIleLysProGlnThrAspArgTrpValPheProAspThrAsnSerGlyI 388  
1169 TCCTCTCTCTGTCGAGGCGCGCTCTTAACTTGGCTGGCTGCTACAGGT 1218  
388 leIleValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 404  
1219 CACCATCTCTGTTATGTCATGTCTATCAAAACAGACACTCGCTCA 1268  
405 HisProSerPheValMetSerCysPheThrAsnGlnValIleAlaG 421  
1269 GCTCGACCTCTACCAAAAAGAGA.....GGAAATCTCGAGAAGAAGGTTT 1312

421 nLeuGluLeuTripAsnGluLysSerSerGlyLysTyrgLulLysValT 438  
1313 ACACACTTCGAGACATCGATGAAGAAGTCGCTCGCTCCACCTCGGA 1362  
438 yrValLeuProProHisLeuAspGluLysValAlaAlaLeuHisLeuGly 454  
1363 TCTCTCGATGTCCACCTTACAAAGCTTACACAGAGCAGGCTGACTACAT 1412  
455 LysLeuGlyAlaLysLeuThrLysLeuSerLysAspGlnAlaAspTyI 471  
1413 CAACGTTCCAGTTCAGGCTCTTACAGCTCTGATGCTTACCGTTAT 1458  
471 eSerValProValGluGlyProTyLysProAlaHisTyIArgTyI 486  
seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAG34060  
seq\_documentation\_block:  
ID AAG34060 standard; Protein; 485 AA.  
XX  
AC AAG34060;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Zea mays protein fragment SEQ ID NO: 41387.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EF1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127482.  
PR 06-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0156559.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

```
alignment_scores:
  Quality: 1523.00      Length: 487
  Ratio: 3.875          Gaps: 4
  Percent Similarity: 80.698  Percent Identity: 61.807

alignment_block:
US-09-759-990-1 x AAG34060 ..

Align seg 1/1 to: AAG34060 from: 1 to: 485

19 ACTGGTGTCTCCATTCGAGTACAGAAATGCGCATCAACCTCCATGTTCT 68
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
8 ThrSerSerGlyArgGluTyrLysValLysAspLeuSerGlnAlaAsp 24
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
69 CGGCCGTAAAGAACTTACCTTCCTGAGAGGAAATGCCAGGCTTATGG 118
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
24 eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuMet 41
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
119 TTCTTCGTGAGCGTTATTCGGTCTTAAAGCCATTGAAGGGTGTGAGAAC 168
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
41 laCysArgAlaGluPheGlyProSerLysProPheAlaGlyAlaArg 57
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
169 TCTGGTTCCTCCATGACATGACATGTCAGACAGCGTCTCATCGAGACACT 218
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
219 CACAGCTCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 268
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
74 uThrAlaLeuGlyAlaGluValAlaArgTrpCysSerCysAsnIlePheSer 91
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
269 CACAAGATACAGCCGTGTGATGATGATGATGATGATGATGATGATGATGAT 318
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
91 hrGlnAspHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 105
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
319 GAGAACCCAGCCGTATCCAGTCTTCGCTGCGCTGGAAGGCGCAACACTGCC 368
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
106 .....ValPheAlaTrpLysGlyGluThrLeuG1 115
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
369 AGAATACCTGGGAAACACATACCGCGCTCTCACATGGCCAGATGGTCAAG 418
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
115 uGluTyrTrpTrpCysThrGluArgCysLeuAspTrpGlyGluAlaGly 132
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
419 GCCCAGCAGAGTTGCGATGATGATGATGATGATGATGATGATGATGATGAT 468
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
132 lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
469 AAGGGCTTC.....GAATTCGAACACGCGGTGTGTTCACGA 506
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
149 GluGlyValLysAlaGluGluGluTyrGluLysThrGlyLysIleProAs 165
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
507 GCCAACAGAGCTGACAACTCGAATACCGCTGCGGTGCTTCTTCTACACACA 556
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
165 pProGluSerThrAspAsnAlaGluPheLysIleValLeuThrIleLea 182
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
557 AGCAGGTCTTCACCAAGAGAAACACCTGGCACACAGTTGCTGCCGCGC 606
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
182 rgAspGlyLeuLysAlaAspProLysLysTyrArgLysMetLysGluArg 198
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
607 ATGAACCGGTGTTCCGAGAGACACAAACAGGTGTCCACCGCTCTACCA 656
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
199 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyrG1 215
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
657 GCTCGAGAGAGGCGGAACTCTCTTCCCGACCATCAACGTCAACGACG 706
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
215 nMetGlnGluThrGlyAlaLeuLeuPheProAlaIleAsnValAsnAsp 232
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
707 CTGTTCACAAAGTCCAGTTCGATACATCTACGGCTGCGCGCACCTCCCTT 756
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
232 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
757 ATCCATGGTATCAACCGGTGCTTCGATGATGATGATGATGATGATGATGATGAT 806
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
```

```
seq_name: /SDSS/gcdata/geneseq/geneseq-emb1/AA2000.DAT:AAG34059
seq_documentation_block:
ID AAG34059 standard; protein; 510 AA.
XX
AC AAG34059;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 41386.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
```

XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 99US-0121825.  
PR 05-MAR-1999; 99US-0121825.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131149.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134769.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 07-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.





```

PR 25-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145931.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 1519.00      Length: 470
    Ratio: 3.966          Gaps: 4
Percent Similarity: 81.489 Percent Identity: 63.404

alignment_block:
US-09-759-990-1 x AAG40087 ..
Align seg 1/1 to: AAG40087 from: 1 to: 467

70 GCGCGTAAGGAACCTTACCCCTGCTGAGAGAAATGCCAGGTCTTATGTT 119
7 GlyArgLeuGluLeuGluAlaGluValGluMetProGlyLeuMetal 23
120 TCTTGTGAGCGTTATTCGGTCTTAAGCCATTGAAGGTTGAGGGTTCAGATCT 169
23 aCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgillet 40
170 CTGGTTCCTCCACATGACAGTCCAGACAGCGTCTCATCGAGACACATC 219
40 hrglySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLeu 56
220 ACAGCTCTGTGTGATGTCAGATGGGCTTCTGCAACATCTTCTCTAC 269
57 ThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerth 73
270 ACAAGATACAGCGGTGCTGCTATCGTTCGCGCCCAACAGGACACACCAG 319
73 rGlnAspHisAlaAlaAlaAlaIleAlaIleAlaIleAlaIleAlaIle 87
320 AGAAGCCAGCGGTATCCAGTCTTCGCTGAGGCGGAAAGGCGGAAACACCTCCA 369
88 .....ValPheAlaTrpLysGlyGluThrLeuGln 97
370 GAATACCTGGAGACACATACCGCTCTCACATGCCAGATGGTCAAGG 419
98 GluTyTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyGln 114
420 CCCACAGCAGGTTCGATGATGGTGTGATGTACACATCTCTCATCTCCA 469
114 yProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHisg 131
470 AGGGCTTCGAA.....TTCGAAACAGCGGTGCTGTTCACAG 507
131 LuGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProAsp 147
508 CCAACAGAGCTGACAACTCGAATACCGTTCGCTTCTTGTACACTCAA 557
148 ProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIleLy 164
558 GCAGGTCTTCAACCAAGAACACACACTGGCACAGATGCTGCTCCCGGCA 607
164 sGluGlyLeuGlnValAspProLysLysTyHisLysMetLysGluArgL 181
608 TGAACGGTGTTCGAGAGACACAAACAGGTGTCACCGCTCTTACAG 657
181 euValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyGln 197
658 CTCGAGAGGAGGCGAAACTCTCTTCCAGCCATCAACGTCAACGAGCGC 707
198 MetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAspSe 214

```



708 TGTTACAAAGTCCAAAGTTCGATACATCTAGGCGTGGCCGACCTCCCTTA 757  
214 rValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeuP 231  
758 TCGATGGTATCAACCGTTCGATGTCATGATCGGCGGCAAGACACGCT 807  
231 roAspGlyLeuMetArgAlaThrAspValMetLeuAlaGlyLysValAla 247  
808 CTCGTCATCGGTTACCGCGGATGCGGAAGGCTGCGCTCAATCCCTCCG 857  
248 ValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetLy 264  
858 TGGCCCAAGCGCTCGGTTATCATCATCAGACGACGACCCCAATCTGCGCTC 907  
264 sThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAlaL 281  
908 TCCAGGCTGCCATGGAAGCTACAGGCTCGCGCATCGAGGAAGTCTGTC 957  
281 euGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVal 297  
958 AAGGATGTCGATATCTTCGTTACATCCACAGGAACTGCGATATCATCTC 1007  
298 SerGluAlaAspIlePheValThrThrThrGlyAsnLysAspIleIleMe 314  
1008 TGTTGACATGATGGCGCATGAAGGATGAAGGCTATTGTCGGTAACATCG 1057  
314 tValAspHisMetArgLysMetLysAsnAsnAlaIleValLysAsnIleG 331  
1058 GCCACTTCGATTAACGAATGTATACAGATGGCTCATGAATACCCAGGC 1107  
331 lyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrProGly 347  
1108 ATCAAGCACATCCCAATCAAGCAGACATACGATGCGGATGGGAATCCGAGA 1157  
348 ValLysArgIleThrIleLysProGlnThrAspArgTyrPheProG 364  
1158 TGCCCAACGCT...ATCCRCCTCTTCTGTCGAGGCGCTCTTCTTAACCTTG 1204  
364 uThrLysAlaGlyIleIleValLeuAlaGluArgLeuMetAsnLeuG 381  
1205 GCTCGGCTACAGGTCACCCATCTTTCGTTATGTCATGATCAATCAACAAC 1254  
381 lyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAsn 397  
1255 CAGACACTCGCTCAGCTCAGCTCTACGAAAGAGA.....GGAATCT 1298  
398 GlnValIleAlaGlnLeuGluLeuThrAsnGluLysAlaSerGlyLys 414  
1299 CGAGAAGAGGTTTACACACTTCCGAAGCATCTCGATGAAGAAGTCGCTC 1348  
414 rGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAlaL 431  
1349 GCCTCCACCTCGGATCTCGATGTCACCTTACAAAGTTACACAGAAG 1398  
431 euLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAsp 447  
1399 CAGGCTGACATCAACAGTTCACGTTGAGGTCCTTACAACTGATGC 1448  
448 GlnSerAspTyrValSerIleProIleGluGlyProTyrLysProProH 464  
1449 TTACGCTTAT 1458  
464 sTyrArgTyr 467

seq\_name: /SID85/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA41407

seq\_documentation\_block:

ID: AA41407 standard; Protein: 485 AA.

XX AC AA41407;

XX DT 18-OCT-2000 (first entry)

XX XX

Arabidopsis thaliana protein fragment SEQ ID NO: 51514.  
Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

23-MAR-1999; 99US-0125788.

25-MAR-1999; 99US-0126264.

29-MAR-1999; 99US-0126785.

01-APR-1999; 99US-0127462.

06-APR-1999; 99US-0128234.

08-APR-1999; 99US-0128714.

16-APR-1999; 99US-0129845.

19-APR-1999; 99US-0130077.

21-APR-1999; 99US-0130449.

23-APR-1999; 99US-0130510.

28-APR-1999; 99US-0130891.

30-APR-1999; 99US-0131449.

30-APR-1999; 99US-0132048.

30-APR-1999; 99US-0132407.

04-MAY-1999; 99US-0132484.

05-MAY-1999; 99US-0132485.

06-MAY-1999; 99US-0132486.

06-MAY-1999; 99US-0132487.

07-MAY-1999; 99US-0132863.

11-MAY-1999; 99US-0134256.

14-MAY-1999; 99US-0134218.

14-MAY-1999; 99US-0134219.

14-MAY-1999; 99US-0134221.

18-MAY-1999; 99US-0134370.

19-MAY-1999; 99US-0134768.

20-MAY-1999; 99US-0134941.

21-MAY-1999; 99US-0135124.

24-MAY-1999; 99US-0135353.

25-MAY-1999; 99US-0135629.

27-MAY-1999; 99US-0136021.

28-MAY-1999; 99US-0136392.

01-JUN-1999; 99US-0136782.

03-JUN-1999; 99US-0137222.

04-JUN-1999; 99US-0137528.

07-JUN-1999; 99US-0137502.

08-JUN-1999; 99US-0137724.

10-JUN-1999; 99US-0138094.

10-JUN-1999; 99US-0138540.

14-JUN-1999; 99US-0138847.

16-JUN-1999; 99US-0139119.

16-JUN-1999; 99US-0139452.

17-JUN-1999; 99US-0139453.

18-JUN-1999; 99US-0139492.

18-JUN-1999; 99US-0139454.

18-JUN-1999; 99US-0139455.

18-JUN-1999; 99US-0139456.

18-JUN-1999; 99US-0139457.

18-JUN-1999; 99US-0139458.

18-JUN-1999; 99US-0139459.

18-JUN-1999; 99US-0139460.

18-JUN-1999; 99US-0139461.

18-JUN-1999; 99US-0139462.

18-JUN-1999; 99US-0139463.

18-JUN-1999; 99US-0139750.

18-JUN-1999; 99US-0139763.

21-JUN-1999; 99US-0139817.

```

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0131930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160788.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 1512.00
    Ratio: 3.887
    Percent Similarity: 79.877
    Percent Identity: 61.396

alignment_block:
US-09-759-990-1 x AAG41407 ..

Align seg 1/1 to: AAG41407 from: 1 to: 485

19 ACTGGTGCCTCATTCGAGTACAGAAATTCGCGACATCAACCTCCATGTTCT 68
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
8 ThrSerGlyArgGluTrpLysValLysAspMetSerGlnAlaAsp 24
CGGCCGTGAAGAACTTACCCTTGCTGAGAGGAATGCCAGGTCTTATGG 118
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
24 eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuVal 41
119 TTCTTCGTGACGGTTATTCGCTTCTTAAGCCATTAAGGGTGTCAAGATC 168
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
41 erCysValThrGluPheGlyProSerGlnProLeuLysGlyAlaArg 57
169 TCTGGTTCCTCCACATGACAGTCCAGACACGCCGCTCTCATCTCGACACT 218
:::|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:

```

58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74  
 219 CACAGCTCTTGCTGATGATGATGGCTTCTGCAACATCTTCTCTA 268  
 74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91  
 269 CACAAGATACACCGCTGCTGCTATCGTTCTGGCCCAACAGGACACCA 318  
 91 hrGlnAspHisAlaAlaAlaIleAlaAlaArgAspSerAlaAla..... 105  
 319 GAGAACCCAGCCGGTATCCAGCTTTCGCTGGAAGGGCGAAACACTCCC 368  
 106 .....ValPheAlaTrpLysGlyGluThrLeuG 115  
 369 AGAATACCTGGGAGAACACATACCGCTCTCACATGGCCAGATGGTCAAG 418  
 115 nGluTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 132  
 419 GCCCAGCAGAGCTTGTGATGATGGTGGTATGCTACACCTCTCATCC 468  
 132 LyrProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148  
 469 AAGGCTTCGAA.....TTCGAACACAGCGGTGCTGTTCCAGA 506  
 149 GluGlyValLysAlaGluIlePheAlaLysAsnGlyThrPheProAs 165  
 507 GCCAACAGACCTGACACCTCGAATACCGCTGCTGCTTGTGCTACACTCA 556  
 165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLe 182  
 557 AGCAGGTCTTCAACCAAGACAAAGCACTGGCACACAGTGTGTCGGCG 606  
 182 ysAspGlyLeuGlnValAspProLysLysTrpHisLysMetLysGluArg 198  
 607 ATGAACGGTGTTCGGAAGACAAACACAGGTGTCACCGCTCTACCA 656  
 199 LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTrpG 215  
 657 GCTCGAAGAGGAGGCAACCTCTTCCAGCCATCAACGTCAACGACG 706  
 215 nMetGlnGluThrGlyAlaLeuLeuPheProAlaIleAsnValAsnAs 232  
 707 CTGTTCAAGTCAAGTTCGATACATCTACGGCTGCGCCACTCCCTT 756  
 232 erValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 248  
 757 ATCGATGATCAACCGCTGCTCGATGTCATGATCGGGCGGCAACAGC 806  
 249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVal 265  
 807 TCTCGTCAGGGTTACGGGATGTCGCAAGGGCTGCGCTCAATCCCTCC 856  
 265 aValIleCysGlyTyrglyAspValGlyLysGlyCysAlaAlaMetL 282  
 857 GTGGCCAGCGCTCGGTTATCATCATCAGAACTCGACCCCAATCTCGCT 906  
 282 ysThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298  
 907 CTCACGGCTGCCATGGAAGGCTACAGGTCGCGCCATCGAGGAGTGT 956  
 299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspVal 315  
 957 CAAGGATGTCGATATCTGTTATCATCAGAGAACTGCGATATCATCT 1006  
 315 lSerGluAlaAspIlePheCysThrThrThrGlyAsnLysAspIleLe 332  
 1007 CTGTGTCATGATGGCCAGATCAAGGATAGGCTATGTGGGTAAACATC 1056  
 332 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348  
 1057 GCCACTTCGATAACGAAATTCATACAGATGGGCTCATGAATACCCAGG 1106  
 349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTrpProG 365

1107 CATCAAGCACATCCCAATCAAGCCAGAAATACGACATGTGGAAATCCCG 1156  
 365 yValLysArgIleThrIleLysProGlnThrAspArgTrpValPhePro 382  
 1157 ATGCCACAGCT...ATCCTCTCTTGTGAGGGCGGCTTCTTAACCTT 1203  
 382 spThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398  
 1204 GCCTGCCCTACAGTCAACCCATCTTCTGTTATGTCAATGTCATTCACAA 1253  
 399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415  
 1254 CCAGACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAATC 1297  
 415 nGlnValIleAlaGlnLeuGluLeuTrpAsnGluLysSerSerGlyLys 432  
 1298 TCAGAGAAGGTTTACACACTTCCGAAGCATCTCGATGAAGAAGTCGCT 1347  
 432 yrGluLysLysValTrpValLeuProLysHisLeuAspGluLysValAla 448  
 1348 CGCTCAGCTCGATCTCTCGATGTCACCTTACAAAGCTTACACAGAA 1397  
 449 AlaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLysAs 465  
 1398 GCAGGCTGATCATCAACCTTCCAGTTCAGGTTCCTTACAGTCTGATG 1447  
 465 pGlnSerAspTrpValSerIleProValGluGlyProTrpLysProValH 482  
 1448 CTTACCGTTAT 1458  
 482 lstrArgTyr 485

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA2000  
 seq\_documentation\_block:  
 ID AA2000 standard; Protein; 497 AA.  
 AC AA2000 standard;  
 DT 18-OCT-2000 (first entry)  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 51513.  
 DE Arabidopsis thaliana  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 OS Arabidopsis thaliana.  
 XX Arabidopsis thaliana.  
 PN EP1033405-A2.  
 PD 06-SEP-2000.  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.

us-09-759-990-1.n2p.rag

Mon Aug 19 15:16:18 2002

PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0132456.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137328.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155559.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.



1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrIle 74

219 CACAGCTCTGTGGTCTGATGTGCAGATGGCTTCCTGCCAACACTTCTCTCA 268  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
74 uThrAlaLeuGlyAlaGluValArgTrpCysSercysAsnIlePheSerT 91

269 CACAAGATACAGCGGCTGCTGCTATCGTTTGTCGGCCCAACAGCACCA 318  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
91 hrGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105

319 GAGAAGCCAGCGGTATCCCAGTCTTCGCCGTGAAGGGCGAAACACTCCC 368  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
106 .....AlaPheAlaTrpLysGlyLeuThrLeuGI 115

369 AGAATACTGGGAGAACACATACCGGGCTCTCATATGGCCAGATGGTCAAG 418  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
115 nGluTyrrTpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 132

419 GCCACACAGAGTGTTCGATGATGGTGTGATGCTACACTCCTCATCTCC 468  
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
132 lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148

469 AAGGGCTTCGAA.....TTCGAAACAGCGGCTGCTGCTCCAGA 506  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
149 GluGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProAs 165

507 GCCAACAGAGCTGACACCTCGATACCGCTGCGTCTTCGTCTACACTCA 556  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIle 182

557 AGCAGGCTTCAACCAGACAGAAGAACCACTGGCACACAGTGTGTCGCGC 606  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
182 ysGluGlyLeuGlnValAspProLysLysTyrrHisLysMetLysGluArg 198

607 ATGAACCGTGTTCGAGAGACAAACACAGGTGCTCCACGCCCTCTACCA 656  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
199 LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTyrgl 215

657 GCTCGAAGAGAGGCGCAACTCCTCTCCAGCCATCAACGTCAACGACG 706  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
215 nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsps 232

707 CTGTTACAAAGTCCAAGTTCGATACATCAGCGTCCGCGCACTCCGCTT 756  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
232 erValThrLysSerLysPheAspAsnLeuTyrrGlySerGlyHisSerLeu 248

757 ATCGATGTATCAACCGTCTCCGATGTATGATCGCGCGCACAGACAGC 806  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
249 ProAspGlyLeuMetArgAlaThrAspValMethIleAlaGlyLysValAl 265

807 TCTGCTCATGGTTACGGCGATCGCGCAAGGCGTGGCGCTCAATCCCCCTC 856  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
265 avallieCySGlyTyrrGlyAspValGlyLysGlyCysAlaAlaMetL 282

857 GTGGCCAAAGCGCTCGCGTTATCATCACAACTCGACCCCAATCTGCGCT 906  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
282 ysthrAlaglyAlaArgValIleValThrGluIleAspProilecysAla 298

907 CTCAGGCTGCATGAAGGCTACCAAGTCCGCGCATCGAGGAAGTCGT 956  
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 315

957 CAAGGATGTCGATATCTCGTTACATGCACAGAACTGCGCATATCATCT 1006  
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
315 lserGluAlaAspilePheValThrThrThrGlyAsnLysaspIleIleM 332

1007 CTGTTGCATGATGCCGCCACATGAAGGATTAAGGCTATTGTTCGTGAACATC 1056  
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
332 etValAspHisMetArgLysMetLysThrAsnProIleValSerThrIle 348

1057 GGCCACTTCGATTAACAAAATTGATACAGATGGCCTCATGAATACCCAGG 1106

349 GHVHSPhesAsnAsnGluIleAspMetProGlyLeuGluThrTyrrProGI 365



```

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 1494.00
    Ratio: 3.952
    Percent Similarity: 80.426
    Percent Identity: 62.766

alignment_block:
US-09-759-990-1 x AAG41408 ..

Align seg 1/1 to: AAG41408 from: 1 to: 467

70 GGCGTAAAGAACTTACCTTGTGAGAAAGAAATGCCAGGTCTTATGTT 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 GlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuValSe 23
120 TCTTCGTGAGCGTATTCCGCTTCTAAGCCATTGAAGGTGTCAAACT 169
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 rCysValThrGluPheGlyProSerGlnProLeuLysGlyAlaArgile 40
170 CTGGTTCCTCCACATGACATCCAGCGCTCTCATCGAGACACATC 219
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
40 hrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLeu 56
220 ACAGCTTTCGTGCTGATGATGATGGCTTCTCTGCAACATCTTCTAC 269
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
57 ThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerTh 73
270 ACAAGATACAGCCGCTGCTGCTATCGTTGTCGGCCCAACAGGCACAC 319
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 rGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 87
320 AGAAGCCAGCCGGTATCCAGTCTTCGCTGGAAGGGGGAACACTCCCA 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 .....ValPheAlaTrpLysGlyGluThrLeuGln 97
370 GAATACTGGGAGAACATACCGGCTCTCATGCGCCAGATGGTCAAGG 419
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GlutyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 114
420 CCCACAGCAGGTTCGATGATGGTGTGATGCTACATCTCTCATCTCCA 469
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 yProAspLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHisG 131
470 AGGCTTCGAA.....TTCGAAACAGCCGGTGTCTGTTCCAGAG 507
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 LuGlyValLysAlaGluGluIlePheAlaLysAsnGlyThrPheProAsp 147
508 CCAACAGAAGCTGACACCTGATACCGCTGCTGCTGCTGCTGCTGCTCA 557
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 ProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIle 164
558 GCAGGTCTTCAACCAAGAACACCACTGGCACACAGTTGCTGCGCGCA 607
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 sAspGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 181
608 TGAACGGTGTTCGGAAGAGACAAACACAGGTGTCCACCGCTCTTACCAG 657

```



```

181 euValGlyValSerGluThrGluThrThrGlyValLysArgLeuTyrGln 197
:: |||||
658 CTCGAGAGGAGGCAAACTCCCTTCCCGAGCCATCAACGTCACGAGCGC 707
::: ||| |||||
198 MetGlnGluThrGlyAlaLeuLeuPheProAlaIleAsnValAspSe 214
708 TGTTACAAAGTCCAACTTCGATCAACATCTACGGCTGCCGCACTCCCTTA 757
::: |||||
214 rValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeuP 231
758 TCGATGGTATCAACCGTCTCGATGTCATGATCGCGGCAACAGACT 807
||| |||||
231 roAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAla 247
808 CTCGTCATGGGTTACGGCGATGTCGCAAGGCTGGCTCAATCCCTCCG 857
::: |||||
248 ValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetLy 264
858 TGGCCAAAGCGCTCGGTTATCATCACAGAACTCGACCCCAATCTGCCCTC 907
::: |||||
264 sThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAlaL 281
908 TCCAGGCTGCCATGGAAGCTACCGAGTCCGCGCATCGAGGAAGTCGTC 957
||| |||||
281 euGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVal 297
958 AAGGATGTCGATATCTTCGTACATGCACAGAACTCGCATATCATCTC 1007
::: |||||
298 SerGluAlaAspIlePheCysThrThrThrGlyAsnLysAspIleIleMe 314
1008 TGTTCACATGATGGCCCGCAGATGAAGGCTAAGGCTATTGTCGGTAAACATCG 1057
||| |||||
314 tValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIleG 331
1058 GCCACTTCGATAAGAAATGTACAGATGGCTCATGAATACCCAGGC 1107
||| |||||
331 lyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrThrProGly 347
1108 ATCAAGCACATCCCAATCAAGCCAGAAATACGACATGTGGGAATTCGCCAGA 1157
::: |||||
348 ValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProAs 364
1158 TGCGCCAGCT...ATCCCTCTCTTGTGAGGCGCGCTCTCTTAACCTTG 1204
| : : : :
364 pThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeuG 381
1205 GCTGCGCTACAGGTACCCATCTTTCGTATGTCATGTCATTCATAAAC 1254
||| |||||
381 lyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAsn 397
1255 CAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAAATCT 1298
||| |||||
398 GlnValIleAlaGlnLeuGluLeuTrpAsnGluLysSerSerGlyLysTy 414
1299 CGAGAAGAGGTTTACACACTTCGAGAGCATCTCGATGAAGAGTCGCTC 1348
||| |||||
414 rGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAlaA 431
1349 GCTCCACCTCGGATCTCTCGATGTCACCTTACAAAGCTTACACAGAG 1398
||| |||||
431 lAlaHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLysAsp 447
1399 CAGGCTGATCATCAACAGTTCAGTTGAGGCTCTTACAGTCTGATGC 1448
||| |||||
448 GlnSerAspTyrValSerIleProValGluGlyProTyrLysProValHi 464
1449 TTACCGTTAT 1458
||| |||||
464 sTyrArgTyr 467

```

seq\_name: /SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:AAW01458

```

seq_documentation_block:
ID AAW01458 standard; Protein: 485 AA.
XX
AC AAW01458;
XX
DT 22-FEB-1997 (first entry)
XX
DE Asparagus S-adenosyl-L-homocysteine hydrolase.
XX
KW S-adenosyl-L-homocysteine hydrolase; SHH; promoter; monocot; dicot;
transgenic plant; disease resistance; pathogen resistance.
XX
OS Asparagus officinalis.
XX
FH Key Location/Qualifiers
FT Region 150..190
FT /note= "region found in SHH of photosynthetic
FT species"
FT Binding-site 263..294
FT /label= NAD+_binding_site
XX
PN WO9632488-A1.
XX
PD 17-OCT-1996.
XX
PF 10-APR-1996; 96WO-CB00882.
XX
PR 10-APR-1995; 95GB-0007381.
XX
PA (ZENE ) ZENECA LTD.
XX
PI Draper J, Greenland AJ, Skipsey M, Warner S;
XX
DR WPI: 1996-477138/47.
XX
DR N-PSDB: AAT44513.
XX
PT S-adenosyl-L-homocysteine hydrolase promoter - used for driving
PT expression of effector genes, such as pathogen resistance genes, in
XX transgenic plants
XX
PS Example 1; Fig 1; 57pp; English.
XX
CC Asparagus S-adenosyl-L-homocysteine hydrolase (SHH) (AAW01458)
CC catalyses the reversible hydrolysis of S-adenosyl-L-homocysteine
CC to adenosine and homocysteine. It contains an extra stretch of
CC amino acid residues (positions 150-190) previously found in other
CC photosynthetic species, parsley and Rhodospirillum rubrum, but not
CC in SHHs from non-photosynthetic species. A cDNA sequence (AAT44513)
CC coding for the asparagus was used to identify the Arabidopsis
CC thaliana SHH gene (AAT44515) and promoter (AAT44514), useful for
CC expression of effector genes in transgenic plants.
XX
SQ Sequence 485 AA;

```

# alignment\_scores:

Quality: 1490.00 Length: 482  
Ratio: 3.821 Gaps: 4  
Percent Similarity: 80.913 Percent Identity: 60.996

## alignment\_block:

US-09-759-990-1 x AAW01458 ..

Align seg 1/1 to: AAW01458 from: 1 to: 485

```

34 GAGTACAGAAATTGCGACATCAACCTCCATGTTCTCGCGCGTAAGAACT 83
|||||
13 GluTyrLysValLysAspMetSerGlnAlaaspPheGlyArgLeuGluI 29
|||||
84 TACCTTCTGTCGAGAAAGAAATGCCAGGTCTTATGTTCTTCGTGAGCGTT 133
: |||||
29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgAlaGluP 46
|||||

```



PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132047.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145513.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.

```
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 1473.00      Length: 459
  Ratio: 3.939          Gaps: 4
  Percent Similarity: 81.481  Percent Identity: 63.181

alignment_block:
US-09-759-990-1 x AAG34061 ..

Align seg 1/1 to: AAG34061 from: 1 to: 450

103 ATGCGAGGCTTATGCTGCTGAGCGTTATTCGCTTCAAGCCATT 152
|||||
1 MetProGlyLeuMetAlaCysArgAlaGluPheGlyProSerLysProPh 17

153 GAAGGGTGCAGAACTCTGTTCCCTCCACATGACAGTCCAGCAGCG 202
: |||||
17 eAlaGlyAlaArgIleSerGlySerLeuHisMetThrIleGlnThrAlav 34

203 TCCTCATCGACACATCACAGCTCTTGGTGTGATGTCAGATGGCTTCC 252
|||||
34 alauilleGluThrLeuThrAlaLeuGlyAlaGluValArgTyrCysSer 50

253 TGCAACATCTTCTACACAAGATACAGCGCTGCTGCTATCGTTGTCGG 302
|||||
51 CysAsnIlePheSerThrGlnAspHisAlaAlaAlaIleAlaArgAs 67

303 CCCAACAGGCACACACAGAGCCAGCGGTATCCAGTCTTCCGCTGGA 352
: |||||
67 pSerAlaAla.....ValPheAlaIleTrpL 75

353 AGGGCGAAACACTCCAGATACTGGGAGAACACATACCGCGCTCTCACA 402
|||||
75 yGlyGluThrLeuGluGluTyrTrpIrpCysThrGluArgCysLeuAsp 91

403 TGCCAGATGGTCAAGGCCACACAGCGTGTGTCATGATGGTGTGATGC 452
||| :|||
92 TrpGlyGluAlaGlyGlyProAspLeuIleValAspAspGlyGlyAspAl 108

453 TACACTCTCATCTCCAAAGGCTTC.....GAATTCCGAACACAG 490
|||||
108 aThrLeuLeuIleHisGluGlyValLysAlaGluGluTyrGluLysT 125

491 CGGTGTGTTTCCAGAGCCACAGAGCTGACACCTCGAATACCGCTGC 540
: ||| :|||
125 hrGlyLysIleProAspProGluSerThrAspAsnAlaGluPheLysIle 141

541 GTTCTTCTACATCAAGAGGCTTCAACCAAGACAGAACCACTGGCA 590
|||||
142 ValLeuThrIleIleArgAspGlyLeuLysAlaAspProLysLysTyrAr 158

591 CACAGTTGCTGCCGCGCATGACGGTGTTCGAGAGACACAACAGGTG 640
: :|||
158 GlyMetLysGluArgLeuValGlyValSerGluGluThrThrGlyL 175
```

seq.name: /SDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAG22957  
seq\_documentation\_block:  
ID AAG22957 standard; Protein; 467 AA.

[illegible]

```

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 31-AUG-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 1457.00
    Ratio: 3.865
    Percent Similarity: 80.901
    Percent Identity: 62.232

alignment_block:
US-09-759-990-1 x AAG22957 ..

Align seg 1/1 to: AAG22957 from: 1 to: 467

19 ACRGGTCTCCATTCGATCAGAAATGCCGACATCAACCTCCATGTTCT 68
   |||:||||| |||:||||| |||:||||| |||:|||||
8 Thr-SerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24

69 CGGCCGTAAGGAACCTTACCCTGCTGAGAGGAATGCCAGTCTTATGG 118
   |||:||||| |||:||||| |||:||||| |||:|||||
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMeCA 41

119 TTCTCTGTCGAGCGTTATTCCTCTTAAGCCATTGAAGGGTGTCAATC 168

```

```

:: ||| ||| :|||:||||| |||:||||| |||:||||| |||:|||||
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57

169 TCTGTTCTCCATCCACATGACAGTCCAGACAGCGCTCTCTCATCGAGACACT 218
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74

219 CACAGCTCTTGTTGTCGATGTCAGATGGGCTTCTGCAACATCTTCTCTA 268
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91

269 CACAAGATACAGCGCTGTCGATGTCGATGTCGCGCCCAACAGGCACACCA 318
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
91 hrGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105

319 GAGAAGCCAGCGGTATCCCGCTTCTGCTGGAAGGCGGAAACACTCCC 368
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
106 .....ValPheAlaTrpLysGlyGluThrLeuG1 115

369 AGAATACTGGGAGAACACATACCGCGCTCTCACATGCCGACATGGTCAAG 418
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
115 nGluTyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGly 132

419 GCCCAGACAGGTTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 468
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
132 lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148

469 AAGGGCTTCGAA.....TTCGAAACAGCGGTGCTGTTCTCCAGA 506
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
149 GluGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProAs 165

507 GCCAACAGAACGTCGACAACTCGAATACCGCTGCTGCTGCTGCTGCTGCT 556
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIleL 182

557 AGCAGGTCCTTCAACCAAGAACCACTGCGCACACAGATTGCTGCCGGC 606
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
182 ySGluGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 198

607 ATGAACGGTGTTCGAGAGACACAACTGCTGCTGCTGCTGCTGCTGCTGCT 656
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
199 LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTyrG1 215

657 GCTCGAAGAGGAGGCAAACTCTCTTCCAGCCATCAACGTCACAGCAGC 706
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
215 nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 232

707 CTGTTACAAAGTCAAAGTTCGATACATCTACGGCTGCGCGCCACTCCCTT 756
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
232 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248

757 ATCGATGTTATCAACCGCTTCCGATGTCGATGTCGATGTCGATGTCGAT 806
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265

807 TCTGCTGATGGTTACGGCGATGTCGCAAGGCTGCTGCTGCTGCTGCTGCT 856
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
265 aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaAlaMetL 282

857 GTGCCCAAGCGCTCGCGTTATCATCATCAGAACTCGACCCCAATCTGGCT 906
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
282 yThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298

907 CTCGAGGCTGCCATGGAAGGCTTACCAGGTCGCGCGCATCGAGGAAGTCGT 956
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 315

957 CAAGGATTCGATATCTTCGTTACATGTCAGAGAACTCGGATATCATCT 1006
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
315 lSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332

1007 CTGTTGACATGATGCCCGCAGATGAAGGATAAGGCTATTGTCGCTAACATC 1056
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

```



```
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146385.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

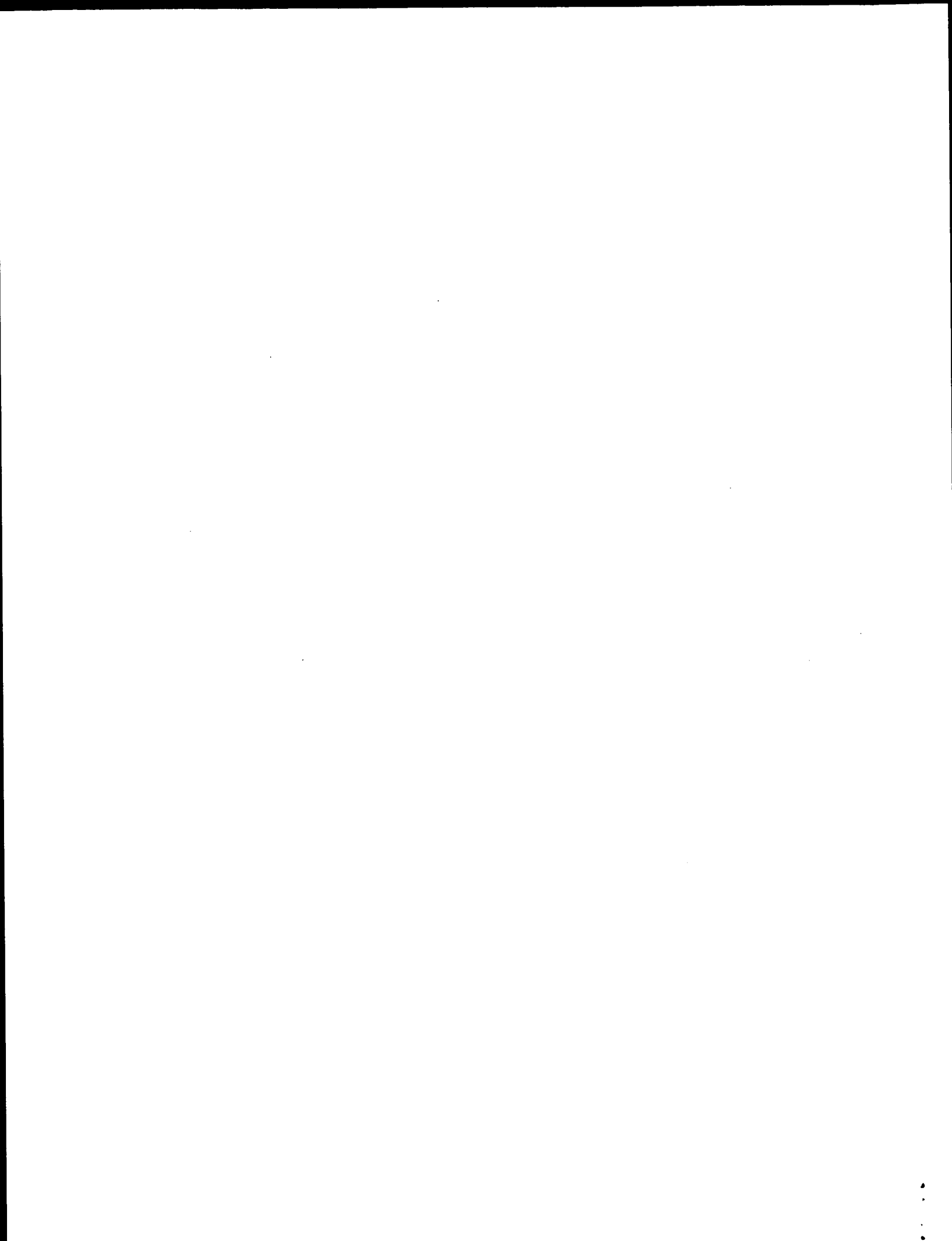
alignment_scores:
    Quality: 1457.00      Length: 466
    Ratio: 3.865          Gaps: 4
Percent Similarity: 80.901 Percent Identity: 62.232

alignment_block:
US-09-759-990-1 x AAG22956 ..
Align seg 1/1 to: AAG22956 from: 1 to: 489

19 ACTGGTCTCCATTTCGAGTACAGAAATTCGCCGACATCAACCTCCATCTTCT 58
   |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
30 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 46
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
69 CGGCCGTAAAGAACTTACCCTTGCTGAGAGGAATGCCAGGTCTTATGG 118
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
46 eGlyArgLeuGluLeuGluLeuAlaGluValGluMetProGlyLeuMeta 63
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
119 TTCTTCGTGAGCGTTATTCGCTTCTTAAGCCATTGAAGGTTCTCAGATC 168
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
63 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile 79
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
169 TCTGGTTCCCTCCACATGACAGTCCAGACAGCCGCTCCTCATCGACACT 218
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
80 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrle 96
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
219 CACAGCTCTTGGTGCTGATGTGATGGGCTTCCTGCAACATCTTCTCTA 268
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
96 uThrAlaLeuGlyAlaGluValArgTyrCysSerCysAsnIlePheSerT 113
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
269 CACAAGATACAGCGCTGCTGCTATGCTGTGCGGCCCAACAGGCACACCA 318
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
113 hrGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 127
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
319 GAGNAGCCAGCCGGTATCCCACTCTCCGCTGGAAGGGGGAACACTCCC 368
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
128 .....ValPheAlaIrrLysGlyGluThrLeuGl 137
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
369 AGAATACCTGGGAGAACATACCGCGCTCTCACATGGCCAGATGGTCAAG 418
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
137 nGluTyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 154
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
419 GCCACAGCAGGTTGTCGATGATGGTGTGATGCTACACTCTCATCTCC 468
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
154 lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 170
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
469 AAGGGCTTCGAA.....TTCGAACAGCCGGTCTGCTTCCAGA 506
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
171 GluGlyValLysAlaGluGluIlePheGluLysThrGlyGluValProAs 187
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
507 GCCAACAGAGCTGACAACTCGAATACCGTGGCTGCTTCTGCTACACTCA 556
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
187 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIleL 204
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
557 AGCAGGCTTCAACCAAGACAGAACCACTGGCACACACTTGTGCGCGGC 606
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
204 ysGluGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 220
   :|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:
```



607	ATGAACCGTGTTTCCGAGAGACACAACAGGTGTCCACCGCCTCTACCA	656
:	:::	
221	LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyrgl	237
657	GCTCGAAGAGGAGGGCAAACCTCTCTCCGCCCATCAAGTCACGACG	706
:	: :::::	
237	nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsps	254
707	CTGTTACAAAGTCCAAGTTCGATAACATCTACGGTGC CGCCACTCCCTT	756
:	:	
254	erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu	270
757	ATCGATGGTATCAACCGTCTCCGATGTCATGATCGCGCGGCAAGACAG	806
:	:	
271	ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValal	287
807	TCTCGTCATGGGTTACGGCGATGTGCGAAGGCTCGCGTCAATCCCTCC	856
:	:	
287	aValIleCysGlyTyrglyAspValGlyLysGlyCysAlaAlaIleMetL	304
857	GTGGCCAAAGCGCTCGGTTATCATCACAGAAGTCCGACCCTCTCGCGT	906
:	:	
304	yThrAlaGlyAlaArgValIleValThrGluIleaspProIleCysAla	320
907	CITCCAGGTCGCATGGAAGCTACAGTCCGCCGATCGAGAGAAGTCGT	956
:	:	
321	LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValva	337
957	CAAGGATGTCGATATCTTCGTTACATGCACAGGAAGTCCGATATCATCT	1006
:	:	
337	IserGluAlaaspliepheValThrThrGlyAsnLysAspIleilem	354
1007	CTGTGACATGATGGCCAGATGAAGATAAGGCTATTTCGCGTAACATC	1056
:	:	
354	etValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle	370
1057	GGCCACTTCGATTAACGAATGTGATACAGATGCCCTCATGAATACCCAG	1106
:	:	
371	GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrrProgl	387
1107	CATCAAGCACATCCCNAATCAAGCCAGAACATGACATGTGGGAATCCCCAG	1156
:	:	
387	yValLysArgIleThrIleLysProGlnThrAspargTrpValPheProg	404
1157	ATGGCCACGCT...ATCCTCCTTCTGCTGAGGGCCGCTCTCTTAACCTT	1203
:	:	
404	LuthrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu	420
1204	GGCTGGCTACAGGTACCCCATCTTTCGTTATGTCATATGTCATCACAAA	1253
:	:	
421	GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs	437
1254	CCAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAAATC	1297
:	:	
437	nGlnValIleAlaGlnLeuGluLeuTrpAsnGlnLysAlaSerGlyLysr	454
1298	TCGAGAAGAAGGTTTACACACTTCCGAAGCATCTCGATGAAGAAGTCGT	1347
:	:	
454	yrGluLysLysValTyrrValLeuProLysHisLeuAspGluLysValAla	470
1348	CGCCTCCACCTCGGATCTCTCGATGTCCTCCACTTACAAAGCTTACACAG	1395
:	:	
471	LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLys	486



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:55:32 ; Search time 80.77 seconds  
(without alignments)  
4443.117 Million cell updates/sec

Title: US-09-759-990-1  
Perfect score: 1461  
Sequence: 1 atggttgcacatcacctac.....ctgatgttaccgttattaa 1461

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2.6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	490	33.5	1812	2	US-08-669-536-1
2	489	33.5	1767	3	US-08-930-894-1
3	375.4	25.7	441529	4	US-09-103-840A-1
4	223.2	15.3	2211	4	US-09-318-448-26
5	122.4	8.4	289	1	US-08-204-740-8
6	122.4	8.4	289	3	US-09-081-167A-8
7	122.4	8.4	289	3	US-09-081-395-8
8	122.4	8.4	289	4	US-09-416-833-8
9	122.4	8.4	289	5	PCT-US95-02521-8
10	105	7.2	285	1	US-08-204-740-6
11	105	7.2	285	3	US-09-081-167A-6
12	105	7.2	285	3	US-09-081-395-6
13	105	7.2	285	4	US-09-416-833-6
14	105	7.2	285	5	PCT-US95-02521-6
15	96.6	6.6	584	4	US-08-998-416-249
16	56.8	3.9	636	4	US-09-328-111-850
17	41	2.8	390	4	US-09-197-649-7
18	40.2	2.8	390	4	US-09-197-649-7
19	40	2.7	2226	2	US-08-896-005-2
20	38.4	2.6	300	2	US-08-716-942-15
21	38.4	2.6	300	2	US-08-716-942-19
22	37	2.5	4403765	4	US-09-103-840A-2
23	36.8	2.5	298	2	US-08-716-942-12
24	36.6	2.5	539	3	US-08-828-741B-12
25	36.6	2.5	539	3	US-09-160-567-12
26	36.6	2.5	599	3	US-08-828-741B-7
27	36.6	2.5	599	4	US-09-160-567-7

28	36.6	2.5	1031	3	US-08-828-741B-5	Sequence 5, Appli
29	36.6	2.5	1031	4	US-09-160-567-5	Sequence 5, Appli
30	36.6	2.5	1490	3	US-08-828-741B-3	Sequence 3, Appli
31	36.6	2.5	1490	4	US-09-160-567-3	Sequence 3, Appli
32	35.8	2.5	1229	1	US-08-440-861-1	Sequence 1, Appli
33	35.8	2.5	1229	1	US-08-433-854-1	Sequence 1, Appli
34	35.8	2.5	1229	1	US-08-174-745A-1	Sequence 1, Appli
35	35.8	2.5	1229	2	US-08-195-947-1	Sequence 1, Appli
36	35.8	2.5	1229	2	US-08-433-885-1	Sequence 1, Appli
37	35.8	2.5	1229	2	US-08-433-908B-1	Sequence 1, Appli
38	35.8	2.5	1229	4	US-08-410-614-1	Sequence 1, Appli
39	35.8	2.5	2000	4	US-09-329-350-34	Sequence 34, Appli
40	35.6	2.4	1155	2	US-08-387-942C-7	Sequence 20, Appli
41	35.4	2.4	1155	2	US-08-387-942C-20	Sequence 20, Appli
42	35.4	2.4	12588	2	US-08-387-942C-1	Sequence 1, Appli
43	35.2	2.4	28804	2	US-08-592-874-1	Sequence 1, Appli
44	35.2	2.4	28804	3	US-09-096-942-2	Sequence 2, Appli
45	35.2	2.4	28804	3	US-09-096-867-2	Sequence 2, Appli

## ALIGNMENTS

## RESULT 1

US-08-669-536-1  
; Sequence 1, Application US/08669536  
; Patent No. 5910444  
; GENERAL INFORMATION:

APPLICANT: MASUTA, CHIKARA  
APPLICANT: UEHARA, KYOKO

APPLICANT: TANAKA, HIDEO  
APPLICANT: KAWATA, SHIGERU

TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH  
STREET: PO BOX 747  
CITY: FALLS CHURCH  
STATE: VA  
COUNTRY: USA  
ZIP: 22040-0747

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/669,536

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: MURPHY JR, GERALD M

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1254-128

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1812 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

US-08-669-536-1

Query Match 33.5%; Score 490; DB 2; Length 1812;  
Best Local Similarity 61.2%; Pred. No. 4.4e-141;  
Matches 887; Conservative 0; Mismatches 515; Indels 48; Gaps 4;  
QY 33 cgagtacagaattgcgcacatcaccttccttcgttcgcgcgtaaggaacttacccttgc 92

```

1152 CCCTGGTGTCAAGAGGATCAACATTAAGCCTAAACCCGACAGATGGGTCTTCCCTGACAC 1211
1161 ccac---gctatcctcttctgagggcgcccttcttaaccttggctgctgacagg 1217
1212 CAACAGTGGCATCATTTGCTTGGCTGAGGGTGGCTCTCATGAACCTTGGGATGTCACAGG 1271
1218 tcacccatcttctgtagtcaatgctcattcacaaacagacacactcgcctcagcagcct 1277
1272 ACACCTAGTGTGATGTCGTGCTGCTTCACTAACCAAGTCATTTGCCCAACTCCAGTT 1331
1278 ct-----acgaaagagagaaatctcgagagaaggtttacacacttccgaagcatct 1331
1332 GTGGAATGAAAGAGCAGTGGGAGATATGAGAAAGAAAGTATGCTTGTGCTTGCCTAACACCT 1391
1332 cgatgaagaagtcgctcgctccactcgcctcgcctcgcctcgcctcgcctcgcctcgcct 1391
1392 CGACGAGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
1392 acagagcagcgtgctgactacatcaacgcttccagttcagttcagttcagttcagttcagtt 1451
1452 GAAGGATCAAGCTGACTACATTAAGCTTCCAGTTGAGGGTCCCTTACAAGCCTGCTCACTA 1511
1452 ccgttattaa 1461
1512 CAGGTACTGA 1521

RESULT 2
US-08-930-894-1
; Sequence 1, Application US/08930894
; Patent No. 6037524
; GENERAL INFORMATION:
; APPLICANT: GREENLAND, Andrew James
; APPLICANT: DRAPER, John
; APPLICANT: SKIPSEY, Marc
; APPLICANT: WARNER, Simon
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,894
; FILING DATE: 09-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00882
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9507381.3
; FILING DATE: 10-APR-1995
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1767 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: SHH GENE FROM ASPARAGUS
; FEATURE:
; NAME/KEY: CDS

```

LOCATION: 26..1483  
OTHER INFORMATION: /codon\_start= 26  
US-08-930-894-1

Query Match 33.5%; Score 489; DB 3; Length 1767;  
Best Local Similarity 61.1%; Pred. No. 8.9e-141;  
Matches 886; Conservative 0; Mismatches 515; Indels 48; Gaps 4;

QY 33 cgagtacagaattgccgacatacaoctccatgttctcgccgtaagaacttacccttgc 92  
DB 61 CGAGTACAAGGTCAAGGACATGCTTCAGGCGGACTTCGGCGCCTCGAGATCAGCTGC 120  
QY 93 tgagaagaaatccaggctcttatgttctctgtgagcgttattccgcttctaagccatt 152  
DB 121 TGAGGTGAGATGCCAGGCTCATGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCT 180  
QY 153 gaagggtgtcagaactctctgttccctccacatgacagtcacagacagccgtctctatcga 212  
DB 181 CAAGGGGCGCAAAATCACTGGATCCCTCCACATGACGATCCAAATGCGCTCCTCATCGA 240  
QY 213 gacactcaagctctgt 272  
DB 241 AACCTTAACCGCTCCGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
QY 273 agatacagccgtctgtctatcgttctcgcccaacagggccacacagagaagccagccg 332  
DB 301 GGACCATGCGCGCTGCCATTCGCGGTGACTCCGC----- 336  
QY 333 tatccagctcttcgttggagggcgaaactcccaactcccaacttactgtggagaaacacatccg 392  
DB 337 ---CTCGGTCTTCGCTTGAAGGTGAGACCTCCAGGAGTACTGGTGTGACCGGAGCG 393  
QY 393 cgcctctcacatggcagatgttcaaggccacacagcaggttctcgatgtgtgtgtgtgtgc 452  
DB 394 TGGCTTCGAGTGGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453  
QY 453 tacactctctatcccaagggtt-----cgaaattcgaaacagccggtgtgt 500  
DB 454 CACTCTCTTGATCCATGAGGGGTGAAGCGCGGAGAGTACGAGAGAGCGGGAAGAT 513  
QY 501 tccagagcaacagaagctgacacactcgaataccgtgtgtgtgtgtgtgtgtgtgtgtgtgt 560  
DB 514 GCGGATCGGCGCTTACCGACAAATGCTGAGTTCAGATCTGCTCAATCATCAGGA 573  
QY 561 ggtcttcaaccaagaagaacacactggcacaggttctgcggcgtgaacgtgtgttc 620  
DB 574 TGGGCTCAAGGTGGACCCACCAGTACAGGAAGATGAAGATAGGATGTGCTGTGTC 633  
QY 621 cgaagagcaacaacaggtgtccacgcctctaccagctcgagagagagggcgaactcct 680  
DB 634 GGAGGAGACCAACCGGGTCAAGAGGCTTACCAGATCGAGCTAACAAATTCCTTCT 693  
QY 681 ctccagcaccatcaacgacagcgtgttacaagagtcgaagtcgaatacactcagc 740  
DB 694 TTTCCCTGGATCAATGTCAATGACTCCCTGACCAAGAGCAAGATTGACATCTGTATGG 753  
QY 741 ctgcgcactcccttatcgtatggtatcaacogtcttcgatgtgtgtgtgtgtgtgtgtgtgt 800  
DB 754 ATGCGGCACCTCTTCCGATGGTCTGATGAGGGCCACTGATGTATGATGTGCTGCA 813  
QY 801 gacagctctgtctatggttaccggtgtcggaaggggtgcgtcgaatccctccgtg 860  
DB 814 GGTGTCATGTTCTCGGTTATGTTGATTCGGAGAGGGGTGCTGCTGCTGCTGCTGCTGCTGCT 873  
QY 861 ccaaggcgtcgtgtatcatcacagaactcgaaccaatctgcgtctccagcgtccat 920  
DB 874 GGTGTTCCCGTGTATGTGACGGAGATGACCCCATCTGCTGCTTCAAGCCCTTAAT 933  
QY 921 ggaaggctaccaggtccgcgcgtcagaggaagtcgtaaggatgtcgatcttcgtttac 980  
DB 934 GGAGGGTCTTCAAGGTCCCTACCCCTCGAGATGTGTCTCAGAGCGCGGATATCTTTGTTC 993

QY 981 atgcacaggaactgcgatatcatctctgttgacatgatggccagatgaaggataaaggc 1040  
DB 994 CACCACCGGTAAACAGGACATCATCTGCTGGACCATGAGGATGAAGAACAATGC 1053  
QY 1041 tatttcggttaacatcgccacttcgataaagaaatgatacagatggcctcatgaata 1100  
DB 1054 CATTTGCTCAACATTTGTTGACAAAGAGATGATGACATGCTAGTTTGGAGACATA 1113  
QY 1101 ccagggcatcaagcacatcccaatcaagccagagaatacagaaatgtgggaattccagatgg 1160  
DB 1114 CCCTGGCATCAAGAGAAATCACCATCAAGCCCACTGACCGGTGGGTCTTCCCTGAAAC 1173  
QY 1161 ccac---gctatccctctctctgtgagggccgctctttaaaccttggctgcgtacagg 1217  
DB 1174 CAACACTGCTATAATTTGTTCTGCTGAGGCGCCACTCATGAACCTTGGGTGTGCCACTGG 1233  
QY 1218 tcaccatcttctgttatgtcaatgtcaattcaacaaacacacactcgtcagctcgacct 1277  
DB 1234 TCACCCAGCTTTGCTATGCTGCTGCTTCTTCCACCAACAGGTGATGCTCAGCTAGATT 1293  
QY 1278 ctacgaaaag-----agaggaatctcgagaagaaggtttacacacttccgaagcatct 1331  
DB 1294 GTGGAATCAGAAGGCAAGCGCAAGTATGAGAAAGAGTTTACGTGCTCCCAAGCATCT 1353  
QY 1332 cgatgaagaagtcgctcgcctccacctcgatctctcgatgtccaccttcaaaagcttac 1391  
DB 1354 TGATGAGAAGTAGCAGCGCTTCACTTGGCAAGCTCGGAGCCAAAGCTTACAAAGCTCAG 1413  
QY 1392 acagaagcaggtgactacatacaagttccagttgaggttccttacaagtctgatgctta 1451  
DB 1414 CCTTCACAGCGGACTACATCAGGCTCCCATCGAGGGTCCCTCAACAGCCACCTCACTA 1473  
QY 1452 ccgttatta 1460  
DB 1474 CAGGTACTA 1482

RESULT 3  
US-09-103-840A-1/C  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103.840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 25.7%; Score 375.4; DB 4; Length 4411529;  
Best Local Similarity 61.3%; Pred. No. 1.4e-103;  
Matches 879; Conservative 0; Mismatches 551; Indels 3; Gaps 1;  
QY 32 tcgagtacagaattgccgacatacaactccatgttcttcggttaaggaactacccttg 91  
DB 3629588 TCGATTTTAAGATCGCGACCTGCTCACTAGCGATTTCGGCCGCAAGAACTCCGGATCG 3629529  
QY 92 ctgagaaggaatgccaggtcttattgttcttcgtgagcgtttattccgtcttctaagccat 151  
DB 3629528 CCGAGCAGGAGATCGCGGCTTGTGCTGCGCGGAGTATGCGGAGGTGCAACCC 3629469





STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-09-081-167A-8

Query Match 8.4%; Score 122.4; DB 3; Length 289;  
Best Local Similarity 71.1%; Pred. No. 3.7e-28;  
Matches 162; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 37 tacagaattgccgacatcaacctccatgttctcgcgcgttaaggaacttacccttgctgag 96  
Db 62 TACAAGTCCGCGACATCGGCTGGCTGGGAGCGCAAGGCCCTGGACATTGCTGAG 121  
QY 97 aaggaatccaggtcttattgttcttcgtgagcgttattccgccttcttaagccattgaag 156  
Db 122 AACGAGATCCGGGCGCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAAGCCACTGAAG 181  
QY 157 ggtgtcagaatctctgttccctccacatgacagtcagtcagacagccgcctccatcagagaca 216  
Db 182 GCGGCCCGCATCGCTGGCTGGCTGCATGACCGTGGAGACGGCGCTCCTCATTGAGACC 241  
QY 217 ctcaagctctgtgctgatgtcagatggggtcttctcgaacatcttc 264  
Db 242 CTCGTACCCCTGGTGTGCTGAGGTGCAGTGGTCCAGCTGCAACATCTTC 289

RESULT 8  
US-09-416-833-8  
; Sequence 8, Application US/09416833  
; Patent No. 6197521  
; GENERAL INFORMATION:  
; APPLICANT: Gudkov, Andrei  
; APPLICANT: Kazarov, Alexander  
; APPLICANT: Mazo, Ilya  
; APPLICANT: Roninson, Igor B  
; TITLE OF INVENTION: Methods for Identifying Genetic  
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant  
; TITLE OF INVENTION: Growth in Cancer Cells  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Alligretti & Witcoff, Ltd.  
; STREET: 10 S. Wacker Drive, Suite 3000  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/416,833  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/204,740  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6197521nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: .93,354-C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; TELEX: 910-221-5317  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 289 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-09-416-833-8

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-09-081-167A-8

Query Match 8.4%; Score 122.4; DB 3; Length 289;  
Best Local Similarity 71.1%; Pred. No. 3.7e-28;  
Matches 162; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 37 tacagaattgccgacatcaacctccatgttctcgcgcgttaaggaacttacccttgctgag 96  
Db 62 TACAAGTCCGCGACATCGGCTGGCTGGGAGCGCAAGGCCCTGGACATTGCTGAG 121  
QY 97 aaggaatccaggtcttattgttcttcgtgagcgttattccgccttcttaagccattgaag 156  
Db 122 AACGAGATCCGGGCGCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAAGCCACTGAAG 181  
QY 157 ggtgtcagaatctctgttccctccacatgacagtcagtcagacagccgcctccatcagagaca 216  
Db 182 GCGGCCCGCATCGCTGGCTGGCTGCATGACCGTGGAGACGGCGCTCCTCATTGAGACC 241  
QY 217 ctcaagctctgtgctgatgtcagatggggtcttctcgaacatcttc 264  
Db 242 CTCGTACCCCTGGTGTGCTGAGGTGCAGTGGTCCAGCTGCAACATCTTC 289

RESULT 7  
US-09-081-395-8  
; Sequence 8, Application US/09081395  
; Patent No. 6083746  
; GENERAL INFORMATION:  
; APPLICANT: Gudkov, Andrei  
; APPLICANT: Kazarov, Alexander  
; APPLICANT: Mazo, Ilya  
; APPLICANT: Roninson, Igor B  
; TITLE OF INVENTION: Methods for Identifying Genetic  
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant  
; TITLE OF INVENTION: Growth in Cancer Cells  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 S. Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/081,395  
; FILING DATE: 18-MAY-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6083746nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 93,354-KK  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 289 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-09-081-395-8



	Best Local Similarity	71.1%;	Pred. No. 3.7e-28;	Mismatches	66;	Indels	0;	Gaps	0;
Qy	37	tacagaattgcgacatcaaacctcatgttcttcggcgttaaggaaccttaccccttgctgag	96						
Db	62	TACAAAGTCGGCGACATCGGCTGGCTGGGGACGCAAGGCCCTGGACATTGCTGAG	121						
Qy	97	aaggaaatgccagagtcttatgtttcttcgttagcgattattccgcttcttaagccattgaag	156						
Db	122	AACGAGATGCCGGGCCTGATGGTATGCGGGAGCGGTACTCGGCCCTCAAGCCACTGAAG	181						
Qy	157	gggtctcagaactcttgttccctccaatcacagtcaccagacgcogtctctatcagagaca	216						
Db	182	GGCGCCGCATCGTGCGTGTCTGTCGCACATGACCGTGGAGACGGCGGTCTCATTTGAGACC	241						
Qy	217	ctcacagctcttggtgtgatgtcagatgggcttctctgcaacatatcttc	264						
Db	242	CTCCTCACCCCTGGGTGTCTGAGGTGCAGTGTGGTCTCAGCTTGCACAATCTTC	289						

```

RESULT      9
PCT-US95-02521-8
; Sequence 8, Application PC/TUS9502521
; GENERAL INFORMATION:
;
; APPLICANT:
;
; TITLE OF INVENTION:  Methods for Identifying Genetic
; TITLE OF INVENTION:  Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION:  Growth in Cancer Cells
; NUMBER OF SEQUENCES:  13
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  PCT/US95/02521
; FILING DATE:
;
; INFORMATION FOR SEQ ID NO:  8:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  289 base pairs
; TYPE:  nucleic acid
; STRANDEDNESS:  single
; TOPOLOGY:  linear
; MOLECULE TYPE:  CDNA
PCT-US95-02521-8

```

Query Match	8.4%;	Score 122.4;	DB 5;	Length 289;
Best Local Similarity	71.1%;	Pred. No. 3.7e-28;		
Matches 162;	Conservative	0;	Mismatches 66;	Indels 0; Gaps
QY	37	tacagaattgccagatacaaacctcatgttctcgccgttaaggaacttacccttgctgag	96	
Db	62	TACAAGATCGCCAGCATCGGCTGGCTGGGGACGCAAGCGCCTGCACATTGCTGAG	121	
QY	97	aaggaaatgccaggttcttatggttcttgtagcgtttattccggtctctaagccattgaag	156	
Ddb	122	AACGAGATGCCGGGCTGATGCGTATGGGGAGCGGTACTCGGCCCTCCAAGCCACTGAAG	181	
QY	157	ggtgtcagaatctctggttccctccacatgacagtccagacagcogtctcatcgagaca	216	
Db	182	GGCGCCGCATCGTGGCTGGCTGTCACATGACCGTGGAGAGCGGCCGTCTCATTTGAGACC	241	
QY	217	ctcacagctcttggtctgatgtcagatgggcttctctgcaacatcttc	264	
Db	242	CTCGTCAACCCCTGGGTGCTGAGGTGCAGTGGTCCAGCTGCAACATCTTC	289	

RESULT 10  
US-08-204-740-6  
; Sequence 6, Application US/08204740  
; Patent No. 5753432  
; GENERAL INFORMATION:

```

1 / APPLICANT: Gudkov, Andrei
2 / APPLICANT: Kazarov, Alexander
3 / APPLICANT: Mazo, Ilya
4 / APPLICANT: Roninson, Igor B
5 / TITLE OF INVENTION: Methods for Identifying Genetic
6 / TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
7 / TITLE OF INVENTION: Growth in Cancer Cells
8 / NUMBER OF SEQUENCES: 13
9 / CORRESPONDENCE ADDRESS:
10 / ADDRESSEE: Allgretti & Witcoff, Ltd.
11 / STREET: 10 S. Wacker Drive, Suite 3000
12 / CITY: Chicago
13 / STATE: Illinois
14 / COUNTRY: USA
15 / ZIP: 60606
16 / COMPUTER READABLE FORM:
17 / MEDIUM TYPE: Floppy disk
18 / COMPUTER: IBM PC compatible
19 / OPERATING SYSTEM: PC-DOS/MS-DOS
20 / SOFTWARE: PatentIn Release #1.0, Version #1.25
21 / CURRENT APPLICATION DATA:
22 / APPLICATION NUMBER: US/08/204,740
23 / FILING DATE: 04-MAR-1994
24 / CLASSIFICATION: 435
25 / ATTORNEY/AGENT INFORMATION:
26 / NAME: No. 5753432ban, Kevin E
27 / REGISTRATION NUMBER: 35,303
28 / REFERENCE/DOCKET NUMBER: 93,354-C
29 / TELECOMMUNICATION INFORMATION:
30 / TELEPHONE: 312-715-1000
31 / TELEFAX: 312-715-1234
32 / TELEX: 910-221-5317
33 / INFORMATION FOR SEQ ID NO: 6:
34 / SEQUENCE CHARACTERISTICS:
35 / LENGTH: 285 base pairs
36 / TYPE: nucleic acid
37 / STRANDEDNESS: single
38 / TOPOLOGY: linear
39 / MOLECULE TYPE: cDNA
40 /
41 /
42 /
43 /
44 /
45 /
46 /
47 /
48 /
49 /
50 /
51 /
52 /
53 /
54 /
55 /
56 /
57 /
58 /
59 /
60 /
61 /
62 /
63 /
64 /
65 /
66 /
67 /
68 /
69 /
70 /
71 /
72 /
73 /
74 /
75 /
76 /
77 /
78 /
79 /
80 /
81 /
82 /
83 /
84 /
85 /
86 /
87 /
88 /
89 /
90 /
91 /
92 /
93 /
94 /
95 /
96 /
97 /
98 /
99 /
100 /
101 /
102 /
103 /
104 /
105 /
106 /
107 /
108 /
109 /
110 /
111 /
112 /
113 /
114 /
115 /
116 /
117 /
118 /
119 /
120 /
121 /
122 /
123 /
124 /
125 /
126 /
127 /
128 /
129 /
130 /
131 /
132 /
133 /
134 /
135 /
136 /
137 /
138 /
139 /
140 /
141 /
142 /
143 /
144 /
145 /
146 /
147 /
148 /
149 /
150 /
151 /
152 /
153 /
154 /
155 /
156 /
157 /
158 /
159 /
160 /
161 /
162 /
163 /
164 /
165 /
166 /
167 /
168 /
169 /
170 /
171 /
172 /
173 /
174 /
175 /
176 /
177 /
178 /
179 /
180 /
181 /
182 /
183 /
184 /
185 /
186 /
187 /
188 /
189 /
190 /
191 /
192 /
193 /
194 /
195 /
196 /
197 /
198 /
199 /
200 /
201 /
202 /
203 /
204 /
205 /
206 /
207 /
208 /
209 /
210 /
211 /
212 /
213 /
214 /
215 /
216 /
217 /
218 /
219 /
220 /
221 /
222 /
223 /
224 /
225 /
226 /
227 /
228 /
229 /
230 /
231 /
232 /
233 /
234 /
235 /
236 /
237 /
238 /
239 /
240 /
241 /
242 /
243 /
244 /
245 /
246 /
247 /
248 /
249 /
250 /
251 /
252 /
253 /
254 /
255 /
256 /
257 /
258 /
259 /
260 /
261 /
262 /
263 /
264 /
265 /
266 /
267 /
268 /
269 /
270 /
271 /
272 /
273 /
274 /
275 /
276 /
277 /
278 /
279 /
280 /
281 /
282 /
283 /
284 /
285 /
286 /
287 /
288 /
289 /
290 /
291 /
292 /
293 /
294 /
295 /
296 /
297 /
298 /
299 /
300 /
301 /
302 /
303 /
304 /
305 /
306 /
307 /
308 /
309 /
310 /
311 /
312 /
313 /
314 /
315 /
316 /
317 /
318 /
319 /
320 /
321 /
322 /
323 /
324 /
325 /
326 /
327 /
328 /
329 /
330 /
331 /
332 /
333 /
334 /
335 /
336 /
337 /
338 /
339 /
340 /
341 /
342 /
343 /
344 /
345 /
346 /
347 /
348 /
349 /
350 /
351 /
352 /
353 /
354 /
355 /
356 /
357 /
358 /
359 /
360 /
361 /
362 /
363 /
364 /
365 /
366 /
367 /
368 /
369 /
370 /
371 /
372 /
373 /
374 /
375 /
376 /
377 /
378 /
379 /
380 /
381 /
382 /
383 /
384 /
385 /
386 /
387 /
388 /
389 /
390 /
391 /
392 /
393 /
394 /
395 /
396 /
397 /
398 /
399 /
400 /
401 /
402 /
403 /
404 /
405 /
406 /
407 /
408 /
409 /
410 /
411 /
412 /
413 /
414 /
415 /
416 /
417 /
418 /
419 /
420 /
421 /
422 /
423 /
424 /
425 /
426 /
427 /
428 /
429 /
430 /
431 /
432 /
433 /
434 /
435 /
436 /
437 /
438 /
439 /
440 /
441 /
442 /
443 /
444 /
445 /
446 /
447 /
448 /
449 /
450 /
451 /
452 /
453 /
454 /
455 /
456 /
457 /
458 /
459 /
460 /
461 /
462 /
463 /
464 /
465 /
466 /
467 /
468 /
469 /
470 /
471 /
472 /
473 /
474 /
475 /
476 /
477 /
478 /
479 /
480 /
481 /
482 /
483 /
484 /
485 /
486 /
487 /
488 /
489 /
490 /
491 /
492 /
493 /
494 /
495 /
496 /
497 /
498 /
499 /
500 /
501 /
502 /
503 /
504 /
505 /
506 /
507 /
508 /
509 /
510 /
511 /
512 /
513 /
514 /
515 /
516 /
517 /
518 /
519 /
520 /
521 /
522 /
523 /
524 /
525 /
526 /
527 /
528 /
529 /
530 /
531 /
532 /
533 /
534 /
535 /
536 /
537 /
538 /
539 /
540 /
541 /
542 /
543 /
544 /
545 /
546 /
547 /
548 /
549 /
550 /
551 /
552 /
553 /
554 /
555 /
556 /
557 /
558 /
559 /
560 /
561 /
562 /
563 /
564 /
565 /
566 /
567 /
568 /
569 /
570 /
571 /
572 /
573 /
574 /
575 /
576 /
577 /
578 /
579 /
580 /
581 /
582 /
583 /
584 /
585 /
586 /
587 /
588 /
589 /
590 /
591 /
592 /
593 /
594 /
595 /
596 /
597 /
598 /
599 /
600 /
601 /
602 /
603 /
604 /
605 /
606 /
607 /
608 /
609 /
610 /
611 /
612 /
613 /
614 /
615 /
616 /
617 /
618 /
619 /
620 /
621 /
622 /
623 /
624 /
625 /
626 /
627 /
628 /
629 /
630 /
631 /
632 /
633 /
634 /
635 /
636 /
637 /
638 /
639 /
640 /
641 /
642 /
643 /
644 /
645 /
646 /
647 /
648 /
649 /
650 /
651 /
652 /
653 /
654 /
655 /
656 /
657 /
658 /
659 /
660 /
661 /
662 /
663 /
664 /
665 /
666 /
667 /
668 /
669 /
670 /
671 /
672 /
673 /
674 /
675 /
676 /
677 /
678 /
679 /
680 /
681 /
682 /
683 /
684 /
685 /
686 /
687 /
688 /
689 /
690 /
691 /
692 /
693 /
694 /
695 /
696 /
697 /
698 /
699 /
700 /
701 /
702 /
703 /
704 /
705 /
706 /
707 /
708 /
709 /
710 /
711 /
712 /
713 /
714 /
715 /
716 /
717 /
718 /
719 /
720 /
721 /
722 /
723 /
724 /
725 /
726 /
727 /
728 /
729 /
730 /
731 /
732 /
733 /
734 /
735 /
736 /
737 /
738 /
739 /
740 /
741 /
742 /
743 /
744 /
745 /
746 /
747 /
748 /
749 /
750 /
751 /
752 /
753 /
754 /
755 /
756 /
757 /
758 /
759 /
760 /
761 /
762 /
763 /
764 /
765 /
766 /
```

Query Match	7.2%	Score 105;	DB 1;	Length 285;	
Best Local Similarity	64.7%	Pred. No. 8.6e-23;			
Matches 156;	Conservative	0;	Mismatches 85;	Indels	0;
QY	24	tgtccattcgagtacagaattgccagatcaaacctccatgtttctcggcogtaaggaaact	83		
Db	45	TGATAAACTCCCTACAAAGTCGCGGACATCGGACTGGCCGCTGGGGAGCAGGAAGGTCT	104		
QY	84	tacccttgcagaaggaatgccaggtctttatggtttcttgtagagcgttattccgcttc	143		
Db	105	GGATATAGCTTGAAATGAGATGCCAGGTTGATGCGCATGCGGGAGATGTACTCAGCCTC	164		
QY	144	taagccattgaagggtgtcagaatctctgtttccctccacatgacagtcacagacgcgt	203		
Db	165	CAAGCCACTGAAGGGTCTCGCATTGTGTGCTCGCTGCGCATGACCGTGGAGACTGTGT	224		
QY	204	ctccatcgagacactcacagctcttggctgctatgctcagatgggcttcctgcacaacttt	263		
Db	225	TCATTTAGACTCTCTGTGCCCTGGTGTGTGAGGGCGCGTGGTCCAGCTGCCAATCTTT	284		
QY	264	c 264			
Db	285	C 285			

RESULT 11  
US-09-081-167A-6  
; Sequence 6, Application US/09081167A  
; Patent No. 6083745  
; GENERAL INFORMATION:  
; APPLICANT: Gudkov, Andrei

```

Query Match      7.2%; Score 105; DB 3; Length 285;
Best Local Similarity 64.7%; Pred. No. 8.6e-23;
Matches 156; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

24  tgcctcattcgagtacagaattgccgacatcaaccccatgtctctcggccgtaagaacct 83
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
45  TGATAAACTGCCCTACAAAGTCGCGGACATCGGACTGGCCGCTGGGGACGGGAAGGCTCT 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84  taccttgcgtgagaaggaaatgccaggtctctatggttcttcctgtagcggttattccgcttc 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 GGATATAGCTGGAATGAGATGCCAGGGTGTGATGCGCATGCGGGAGATGATACFCAGCCCTC 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144  taagcattgaagggtgtcagaatctctggttcccatgcacatgacagtcacagacgcgt 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 CAAGCCACTGAAGGGTGCTCCATGTCTGGCTGGCCATGACCGCTGGAGACTGCTGT 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
204  cctcatgagacactcacagctctctggtgctgagtctcagatgggcttcctgcacaactctt 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
225 TCTCATTTGAGACTCTCGTGGCCCTGGGTGCTGAGGCGCGGTGGTCCAGGTGCAACATCTT 284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      264 c 264
Db       285 C 285

RESULT 13
US-09-416-833-6
; Sequence 6, Application US/09416833
; Patent No. 6197521
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya

```

```

Query Match      7.2%; Score 105; DB 3; Length 285;
Best Local Similarity 64.7%; Pred. No. 8.6e-23;
Matches 156; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY      24  tgcctcattcgagtacagaaattgcccgacatcaacctccattgttctcggccgtaagggaact 83
DB      45  TGATAAACTGCCCTACAAAATGTCGGGACATCGGACATGGCGCGCTGGGGACGGAAGGCTCT 104

QY      84  taaccttgctgagaagaaatgccaggctottatggttcttcgtgagcgttatccgccttc 143
DB      105  GGATATAGCTGGAAGTGAATGAGATGCCAGGTTGTATCGCATCGGAGATGTACTCAGGCTC 164

QY      144  taagccattgaagggtgtcagaatctctggtttccctccacatgacagtcacagacagcgt 203
DB      165  CAAGCCACTGAAGGGTGCTCGCATTTGCTGGCTGCTCGCATGACCGTGGGAGACTGGTGT 224

QY      204  ctcattgagaactcacagctcttggtgctgatgctcagatgggcttctctgcaacatctt 263
DB      225  TCATATTGAGACTCTCTGTCCTTGGTGTGTGAGCGCGGTGGTCCAGCTGCAACATCTT 284

QY      264  c 264
DB      285  C 285

RESULT 12
US-09-081-395-6
; Sequence 6, Application US/09081395
; Patent No. 6083746
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander

```

```

; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; NUMBER OF SEQUENCES: 13
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/416,833
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/204,740
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6197521nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-416-833-6

```

```

Query Match 7.2%; Score 105; DB 4; Length 285;
Best Local Similarity 64.7%; Pred. No. 8.6e-23;
Matches 156; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

```

QY 24 tgcctcattgagtagacagaattgcccagacatacaacctcattcttcgcccgaaggaact 83
|| || || || || || || || || || || || || || || || || || || || || ||
Db 45 TGATAAACTGCCCTACAAAGTCGGGACATCGGACCTGGCGCTGGGAGCGAAGGCTCT 104

QY 84 tacccttgctgagaagaaatgccaggtcttattgttcttcgtgagcgttattccgcttc 143
|| || || || || || || || || || || || || || || || || || || || || ||
Db 105 GGATATAGCTGAGATGAGATGCCAGGTTGATCGCATCGGGAGATGTACTCAGCCTC 164

QY 144 taagccattgaagggtgctgcgacttctggttcctccacatgacagtcacagacgcgt 203
||||| ||||| || || || || || || || || || || || || || || || || ||
Db 165 CAAGCCACTGAAGGCTGCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224

QY 204 cctcatcgagacactcacagctcttctggtgctgagtgatgagtggttcctcgaacattt 263
||||| ||||| || || || || || || || || || || || || || || || || ||
Db 225 TCTCATTTAGACTCTCGTGGCCCTGGTGTGAGGCGGCTGGTGTGAGGCGGCTGCAACATCTT 284

QY 264 c 264
Db 285 C 285

```

```

RESULT 14
PCT-US95-02521-6
; Sequence 6, Application PC/TUS9502521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods for Identifying Genetic

```

```

; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02521
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US95-02521-6

```

```

Query Match 7.2%; Score 105; DB 5; Length 285;
Best Local Similarity 64.7%; Pred. No. 8.6e-23;
Matches 156; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

```

QY 24 tgcctcattgagtagacagaattgcccagacatacaacctcattcttcgcccgaaggaact 83
|| || || || || || || || || || || || || || || || || || || || || ||
Db 45 TGATAAACTGCCCTACAAAGTCGGGACATCGGACCTGGCGCTGGGAGCGAAGGCTCT 104

QY 84 tacccttgctgagaagaaatgccaggtcttattgttcttcgtgagcgttattccgcttc 143
|| || || || || || || || || || || || || || || || || || || || || ||
Db 105 GGATATAGCTGAGATGAGATGCCAGGTTGATCGCATCGGGAGATGTACTCAGCCTC 164

QY 144 taagccattgaagggtgctgcgacttctggttcctccacatgacagtcacagacgcgt 203
||||| ||||| || || || || || || || || || || || || || || || || ||
Db 165 CAAGCCACTGAAGGCTGCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224

QY 204 cctcatcgagacactcacagctcttctggtgctgagtgatgagtggttcctcgaacattt 263
||||| ||||| || || || || || || || || || || || || || || || || ||
Db 225 TCTCATTTAGACTCTCGTGGCCCTGGTGTGAGGCGGCTGGTGTGAGGCGGCTGCAACATCTT 284

QY 264 c 264
Db 285 C 285

```

```

RESULT 15
US-09-998-416-249
; Sequence 249, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebeschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwalls Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

Mon Aug 19 15:16:24 2002

```

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/998,416
  FILING DATE: 24-DEC-1997
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: CH 0016/97
    FILING DATE: 31-DEC-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: Meigs, J. Timothy
    REGISTRATION NUMBER: 38,241
    REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 919-541-8587
    TELEFAX: 919-541-8689
    INFORMATION FOR SEQ ID NO: 249:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 584 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
      MOLECULE TYPE: DNA (genomic)
      ORIGINAL SOURCE:
        ORGANISM: PAG1221UP
        US-08-998-416-249

```

[illegible]

Search completed: August 17, 2002, 05:49:35  
Job time: 14043 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:55:57 ; Search time 1886.52 Seconds  
(without alignments)  
10452.607 Million cell updates/sec

Title: US-09-759-990-1  
Perfect score: 1461  
Sequence: 1 atggttgcaaatcaactac.....ctgatgtaccgttattaa 1461

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estio:\*
- 8: em\_hic:\*
- 9: gb\_esti:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrc:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442.2	30.3	1550	10	BM321022
2	343	23.5	1076	10	BG337961
3	339.4	23.2	776	10	BM321423
4	338.8	23.2	843	10	BF256594
5	338.2	23.1	951	10	BM321451
6	333.6	22.8	800	10	BI929547
7	330.2	22.6	977	9	AL163774
8	329.2	22.5	1163	10	BI244550
9	316.2	21.6	908	10	BF263419
10	313.6	21.5	867	10	BF253807
11	311	21.3	809	10	BE559309
12	307.4	21.0	804	10	BM411809
13	303.4	20.8	916	10	BG444868
14	302.8	20.7	769	10	BF863943
15	301.2	20.6	844	10	BF256768
16	300	20.5	730	10	BM112072
17	299.8	20.5	802	12	BH604363

c	18	299.6	20.5	732	10	BI750640
	19	299.2	20.5	862	10	BM321023
	20	299	20.5	726	10	BE413441
	21	298.8	20.5	689	9	AL503944
	22	297.2	20.3	932	10	BG366184
c	23	296	20.3	902	12	AZ679560
	24	295	20.2	1004	9	AL518398
	25	292.6	20.0	628	10	BI874978
	26	292.2	20.0	738	10	BI405371
	27	291.8	20.0	1070	9	BE033415
	28	291.4	19.9	906	10	BF258750
	29	290	19.8	676	9	AV916518
	30	289.2	19.8	774	10	BG599587
	31	288.8	19.8	904	10	BG444506
	32	288.8	19.8	672	10	BG599809
c	33	288.4	19.7	938	12	BH160428
	34	287.6	19.7	1050	9	BE036599
	35	286.4	19.6	893	10	BG478112
	36	286	19.6	908	10	BM453763
	37	285.8	19.6	648	10	BG274691
c	38	285.6	19.5	707	10	BI751148
	39	284.8	19.5	657	10	BI959820
	40	283.8	19.4	607	10	BF253757
	41	283	19.4	741	10	BM412612
	42	282.4	19.3	957	10	BF268594
	43	282	19.3	821	10	BF261131
	44	281	19.2	625	10	BM323515
	45	280.2	19.2	828	10	BI259193

## ALIGNMENTS

## RESULT 1

## LOCUS

## DEFINITION

BM321022 1550 bp mRNA linear EST 03-JAN-2002  
rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library  
Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC 3.3.1.1), mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## BASE COUNT

## ORIGIN

## Query Match

## Best Local Similarity

## Matches 721;

## Conservative

## 0;

## Mismatches

## 407;

## Indels

## 10;

## Gaps

## 3;

QY 324 gccagccggtatccagcttcgctgaagggcgaacacactccccaataactggagaa 383  
Db 44 GCAGCGCGCGCTCTCGGCTTCGCTTGAAGGGCGAGAACCTCCAGAGTACTGGAGTG 103  
QY 384 cacataccgctctcacatggccagatggttcaaggccacagcaggttctgtgatg 443  
Db 104 CACCTGGGAAGCCCTGTCTGGGCCCCCTACCAAGGGCCCTCAGATCATCTCGACGACGG 163  
QY 444 tgggtatgctacactcctcatcctcaagggcttgaattcgaacagcgggtgtcttc 503  
Db 164 CGGTGACGCGACTCTGTATGATCACAAGGGGTTCGGCGCGAGGACAACCCCAAGCTGCT 223  
QY 504 agagcaacagaagctgacaactcctgaatcacgctgtcttctgtctacactcaagcagt 563  
Db 224 GGAG---GACACGAGGGCCCTCAGAGGCTGCCCTGCCCTCAACAACGCTGCTCAAGCAGGT 280  
QY 564 cttaaccaagaacaaacactcgtgcacacagttgtgtcgcggatgaacggtgtttccga 623  
Db 281 CCAGGAAGGAGGAGCCCGGCTTCTGGCACAAGATCTCTCCCGAGATCGCGGCTGTACGCGA 340  
QY 624 agagacaacacaggtgtccacgctctacagctcgagagagggcgaactcctctt 683  
Db 341 GGAGACGACGACTCGGCTGTATGAGGTGTACAGCTGTCCCGGACGACGAGCTGTCTGT 400  
QY 684 ccagccatcaacgtcaacgacgctgttacaagttccaaagtccaaatcctacgtcgtg 743  
Db 401 CCGGCGCTCAACGCTCAAGGACTCTGTCCACCAAGAGCAAGTTTGACAACATCTACGCGTG 460  
QY 744 ccgcaactccttctcgtatggtatcaacacgctgtcctcgtatgctatgctcgcggcagac 803  
Db 461 CCGCCACTCGCTCATCGACGGCATCAAGCGCGGACGACGAGTGTCTCGCGGCAAGGT 520  
QY 804 agctctcgtatggttacgagctgtcgaagggctgcgctcgaatccctccgtgcca 863  
Db 521 CGCGCTGTGCGGGCTACGGGACGCTGGCAAGGGCTGGCCGAGTCTGTGCGCGGCA 580  
QY 864 agggctcgtgtatcatcacagaactcgaacccaatctgcgtctcctcagagtgctcagtgga 923  
Db 581 GGGCTGCGCGCTCATCGACGAGATCGACCCCATCTGCGCGCTGCGAGGCTCGATGGC 640  
QY 924 aggtacacaggtccgccgcatgaggaagtcgtcaagagtgctgatatcttcgtatcg 983  
Db 641 GGGCTGCGAGGTCACACGCTCGAGCGGGGCTCGACCGCGCGAGATCTTGTGTACGCG 700  
QY 984 cacaggaactcgtatcatctctgtgacatgagtcgacagatgaggaataaggctat 1043  
Db 701 GACGGGCAACCAACATCATATGCGCGAGCATGAGCAAGATGCGCCACAACCTCGGT 760  
QY 1044 tgcgtgtaacatcgccacttcgatcaacgaattgtatcagagatggcctcagaaataccc 1103  
Db 761 CGTCTGCAACATCGGCGCACTTTGACAACGAGATCGAGTCTCGCGGGGCTCAGGAAGTGGCC 820  
QY 1104 aggcataagaacatcccaatcaagcagaatacagcatgtgggaattccagatggcca 1163  
Db 821 CGGCGTCAAGTGGCTCAACATCAAGCGCGAGCTGAGTACGATCTTTCGCGAGGCGCA 880  
QY 1164 cgtatcctcctcttctgagggcgcccttcttaaccttggctgagcagatcagtcaccc 1223  
Db 881 CGGCATCATCTCTGCTCGGAGGGCGGCTGTGTCACACNTCG-TGCGGCAAGGGCGCACCC 939  
QY 1224 atcttctgtagtgaatgcatcacaacacagacactcgtcagctcagctcagcttaaga 1283  
Db 940 CAGNTTGTGATGNNNACAGCTTCACAAACAGACGCTGGCCCGAGNTTAAAGTTCGT--- 996  
QY 1284 aaagagagaatctcgaagaaggtttacacacttccgaagcatctcgtatgaagaagt 1343  
Db 997 ---GCCGAGAGAGTGGACATCGCCGTTCATCACCATCCCAAGATCTTCGAGAGAGGT 1053  
QY 1344 cgtcgcctccacactcgtatctcgtatgtctcgaacttacaagaagttacacagaagcaggc 1403  
Db 1054 CGCGCGCTGCACTCGAAGGCTCAAGGTCGTCTNTACCGCGCTGACCCGCGCCAGGC 1113

QY 1404 tgactacatcaacgttccagttgagggctcctacaaagtctgagcttaccggtattaa 1461  
Db 1114 CAAGTACATCGAGTCCCGCTGACGCGCGGTACAAAGCGCGCCACACTACCGCTANTAA 1171

RESULT 2  
BG837961 1076 bp mRNA linear EST 25-MAY-2001  
LOCUS Zm10\_03e04\_A Zm10\_AAPC\_ECORC\_Fusarium\_graminearum\_corn\_silk Zea  
DEFINITION mays cDNA clone Zm10\_03e04, mRNA sequence.  
ACCESSION BG837961  
VERSION BG837961.1 GI:14204284  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 1076)  
AUTHORS Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De  
Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A., Sprott  
D. and Tinker,N.A.  
TITLE Expressed Sequence Tags from Maize Silk Six Hours After Silk  
Channel Inoculation with Fusarium graminearum  
JOURNAL Unpublished (2001)  
COMMENT Contact: Harris, Linda J.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-food Canada  
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
CANADA  
Tel: (613) 759-1314  
Fax: (613) 759-6566  
Email: harris.lj@em.agr.ca.  
FEATURES  
source  
1..1076  
Location/Qualifiers  
/organism="Zea mays"  
/cultivar="CO388"  
/db\_xref="taxon:4577"  
/clone="Zm10\_03e04"  
/clone\_lib="Zm10\_AAPC\_ECORC\_Fusarium\_graminearum\_corn\_silk"  
/tissue\_type="Silk"  
/dev\_stage="4-5 days post-silk emergence"  
/note="Vector: Bluescript SK+/XhoI-ECORI; Site 1: ECORI;  
Site 2: XhoI; Field-grown corn was silk channel-inoculated  
in the morning (~10 am) with 1 ml of a macroconidial  
suspension (500,000 spores/ml) of Fusarium graminearum and  
silk channels were collected and immediately frozen in  
liquid nitrogen 6 hours later. RNA was extracted from  
silk tissue between 1 cm below and above the inoculation  
point in the silk channel. RNA from five silk channels was  
pooled."

BASE COUNT 228 a 281 c 288 g 278 t 1 others  
ORIGIN

Query Match 23.58; Score 343; DB 10; Length 1076;  
Best Local Similarity 67.2%; Pred. No. 9.2e-80;  
Matches 533; Conservative 0; Mismatches 250; Indels 10; Gaps 3;

QY 678 cctcttccagcagcatcaacgtcgaacgctgttcaaaagtccaagttcgtatcaatctta 737  
Db 1 CCCTTCCTCCCTGCCATTAAAGCTCAAGGATTCGTCACCAAGAGCAAGTTTGACAACCTGTA 60  
QY 738 cggctgcgcgaactcccttctcgtatggtatcaacacgctgtcctcgtatgctcagtcggcg 797  
Db 61 TGGTTGCGCGCCTCACTCCCTGTATGGTCTGATGAGGCCACCGAGCTTATGATCGCCGG 120  
QY 798 caagacagctctcgtcgttaccgagtgctcgaagggctgcgctcaatccctccg 857  
Db 121 TAAGTTGCCGTGCTGCGGATACGAGTGTGGCAAGGGTTGGCGCGCTGCACTCAA 180  
QY 858 tg-gccaagggcgtcgtgttatcatcacagaactcgcacccaatctcgtcctccaggtg 916







1230 York Avenue, New York, NY 10021, USA  
Email: mmuller@rockvax.rockefeller.edu  
Insert Length: 951 Std Error: 0.00  
POLYA-No.

## FEATURES

source Location/Qualifiers  
1. 951  
/organism="Mastigamoeba balamuthi"  
/strain="ATCC 30984"  
/db\_xref="taxon:108607"  
/clone\_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
/note="syn: Phreatamoeba balamuthi"  
BASE COUNT 186 a 321 c 303 g 139 t 2 others  
ORIGIN

Query Match 22.1%; Score 338.2; DB 10; Length 951;  
Best Local Similarity 62.2%; Pred. No. 1.6e-78;  
Matches 598; Conservative 0; Mismatches 330; Indels 33; Gaps 3;  
QY 33 cgaatcagaattgcccacatcaactccatgtttctcgccgttaaggaacttacccttgc 92  
DB 24 CGACTCAAGGTCAAGGACATCAACCTCGCGACTTCGGCCGCAAGGAGATCGAGCTGC 83  
QY 93 tgagaagaaatgccaggtcttatgttcttcgtgagcgttattcgcgttcttaagccatt 152  
DB 84 CGAGTACGAGATGCCGCTCTGATGGCCACCAG---GACCAAGTACGGCCCGAGTGCCT 140  
QY 153 gaaggtgtcagaatctctgttctccctccacatgacagtcacagcagcgtctctatcga 212  
DB 141 CCAGGGCGGAGGATACCGGCTCGCTCCACATGACGATCCAGACCGCGCTCTCATCGA 200  
QY 213 gacactcacagcttctgtctgtatgacagtggtcttccttcgcaacatcttcttacaca 272  
DB 201 GACGCTCAAGTCTCTCGGCGCAAGATCCGCTGGTGTCTGGTCAACATCTTCTCGACGA 260  
QY 273 agatacagcgcgtgtctgtatcgttctgcccccaacagggcacacagagaagccagccg 332  
DB 261 GGACACGCGCGCGCGCCCATCG-----CCAGGCCCG 293  
QY 333 tatccagctctcgcctggaagggcgaaacactcccagaatactggggagacacataccg 392  
DB 294 CGTCTCGGTCTTCCCTTGAAGGGCGAGAACCTCCAGGAGTACTGGGAGTGCACCTGGAA 353  
QY 393 cgtctcacatggccagatggtcaaggccacacagcaggttgcgtatgattggtgatgc 452  
DB 354 GGCCTGTGCTTCCGCGCCCTACACAGGCGCTTCAGATCATCTCTCAGCAGCGCGGTGACGC 413  
QY 453 tacactcctcatctcaagggttctgaattcgaacacagcgggtgtgttccagagccaac 512  
DB 414 GACTCTGATGATCCACAAGGGGTTCGCGCGCGAGGACAAACCCCAAGCTGTGGAG---GA 470  
QY 513 agaagctgacaacctcgaataaccgctgcgttcttctgacactcaagcaggtcttcaaca 572  
DB 471 CCAGCAGGCGCTCGAGAGGTGCGCTTGCCTTACACACGTCGTCGAAGCAGGTCCTCAAGA 530  
QY 573 agacaagaaccactggcacacagttgctgcggcgaatgaacggtgtttccggaagacaac 632  
DB 531 GCAGCCGCGCTTCTGGCACAAGATCTCCCGAGATCCCGGCTGTCTCAGCAGGAGACAC 590  
QY 633 aacaggttccaccgcttaccagctcgaagaaggagggaacactctcttccagccat 692  
DB 591 GACTGGCGTATGAGGCTGTACAGCTGACCGCGAGCGGAAGCTGTGTTCGCGCGCT 650  
QY 693 caacgtcaacgagcgtgttacaagttccagttcgtataacatctacggtcgcgcacactc 752  
DB 651 CAACGTCAGGACTCTNTCACCAGAGCAGATTCGACAAATCTTACCGCTGCGCCGCACTC 710  
QY 753 ccttatcagtgatataaacgctgtctcccgatgtcatgtatgctggggggaagacagctctcgt 812  
DB 711 GCTCATACGCGCATCAAGCGGGGACCGACGCTGATGCTCGCGGCAAGGTTCGCGCTCGT 770  
QY 813 catgggttacggcgtatcgcaagggtcgctcctaatcctcctcgttgcgaagcgctcg 872  
DB 872

Db 771 CGCGGGCTAGCGGACGCTGGCGAAGGGCTGCGCGAGTCTGCTGCGCGGCGGAGGCTGCGG 830  
QY 873 cgttatcacaagaactcgaacccaatctcgtctcctcaggtcgtccatgggaaggttacca 932  
DB 831 CBTATCGTACGAGGAGATCGACCCCATCTGCGCGCTGCAAGCGCTGATGCGCGGCTTCA 890  
QY 933 ggtccgcccgcagcaggaagtcgtcaagatgctcagatcttcgtttacatcacacagaaa 992  
DB 891 GGTCAACGCTCGAGCGGGGCTCGACCGCGCGAGATCTTGTGACGGGACGGGCAA 950  
QY 993 c 993  
DB 951 c 951

RESULT 6  
BI929547  
LOCUS  
DEFINITION EST549436 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA  
clone cTOB29E10 5' end, mRNA sequence.  
ACCESSION BI929547  
VERSION BI929547.1 GI:16243679  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 800)  
AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,  
Utterback, T., Van Aken, S., Ronning, C.M., Nierman, W., Fraser, C.M.,  
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
TITLE Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)  
JOURNAL Unpublished (2001)  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3.  
Location/Qualifiers  
1. 800  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone\_lib="tomato flower, 3 - 8 mm buds"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/note="vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Cornell University; sequencing: The  
Institute for Genomic Research. Flower buds and flowers  
were taken from greenhouse plants (4-8 wks old, TA496).  
They were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."

## FEATURES

source  
1. 800  
/organism="Lycopersicon esculentum"  
/db\_xref="taxon:4081"  
/clone\_lib="tomato flower, 3 - 8 mm buds"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/note="vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Cornell University; sequencing: The  
Institute for Genomic Research. Flower buds and flowers  
were taken from greenhouse plants (4-8 wks old, TA496).  
They were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."  
BASE COUNT 200 a 182 c 198 g 220 t  
ORIGIN

Query Match 22.8%; Score 333.6; DB 10; Length 800;  
Best Local Similarity 65.6%; Pred. No. 2.4e-77;  
Matches 521; Conservative 0; Mismatches 264; Indels 9; Gaps 2;

QY 677 tctcttcccagccatcaagctcaagcgtgtgttacaagtcacaggttcgataacatc 736  
DB 5 TGTCTTTCCTGCTATTATGCTTAATGACTCTGTACCAAGAGCAAGTTTGACACTTGT 64  
QY 737 acgggtcgccgactcccttcatcgtatcgtatcaacgctgctccgattcgtatcgagcg 796  
DB 65 ACGGATGCCGCCACTCACTTCCCGATGGTCTCATGAGGGCTACTGATGATTGCTG 124

QY	797	gcaagacagctctcgtcgtatggttaagcgatgctcgcaagggtggtcgaatccctcc	856
Db	125	gaaagctggtcgttctgctggttgaagatgctcgcaagggtggtcgaatccctcc	184
QY	857	gtggcaagcgctcgtcgtatggttaagcgatgctcgcaagggtggtcgaatccctcc	916
Db	185	acaagctggtcgtcgtatggttaagcgatgctcgcaagggtggtcgaatccctcc	244
QY	917	ccatgaagctcaccaggtcgcgcgacaggaagtgctcaaggatgctgatatctgg	976
Db	245	ccatgaagctcaccaggtcgcgcgacaggaagtgctcaaggatgctgatatctgg	304
QY	977	ttacatgacagaaactcgatatactctgttgacatggtggtcgaatccctcc	1036
Db	305	tgaccacacacatggttaacaaagacatcattcattggttgacacatggtggtcgaatccctcc	364
QY	1037	aggtattgctggttaacatggtcgcgcgacaggaatgatacagatggtcctcatga	1096
Db	365	atgccatgctgcaacatggttgacatggttgacacatggttgacacatggttgacac	424
QY	1097	aataccaggtcacaacacatcccaatcaagcagaatacagacatggtggaattccca	1156
Db	425	ccctgctggttggaagaggtacacatcaagcctcaaacccagacatggtggttccca	484
QY	1157	atggcaagctatctctctt---gctgagggcgctcttcttaacctggtcgctga	1213
Db	485	acaccacagtggtcattggttggtggtggtggtggtggtggtggtggtggtggtg	544
QY	1214	cagtcacacatcttctggtatgctcattcattcacaacacacacacacacacacac	1273
Db	545	caggac	604
QY	1274	acctct-----acgaagagaggaatactcgaagaagtttacacacacacacacac	1327
Db	605	agttgtggaatgagaagagcagtggttaatacagagaagagagagagagagagagag	664
QY	1328	atctcgtatgaagagagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1387
Db	665	acctgacag	724
QY	1388	ttacacagaagcaggtgactacatcaacagtcaggttccaggttccaggttccaggtt	1447
Db	725	tttaccagaggtacagctgactacatgaggttaccaggttaccaggttaccaggtt	784
QY	1448	cttaccgttattaa	1461
Db	785	actacaggtactga	798
RESULT 7			
LOCUS	AL163774	977 bp	linear
DEFINITION	AL163774 Scherffellia dubia pSPORT1 Scherffellia dubia cDNA clone 418	EST 04-APR-2001	
ACCESSION	AL163774		
VERSION	AL163774.1	GI:13547742	
KEYWORDS	EST.		
SOURCE	Scherffellia dubia.		
ORGANISM	Scherffellia dubia		
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;		
AUTHORS	Chlorodendreales; Chlorodendraceae; Scherffellia.		
TITLE	Becker, B., Feja, N. and Melkonian, M.		
JOURNAL	Analysis of expressed sequence tags (ESTs) from the scaly green		
MEDLINE	flagellate Scherffellia dubia Pascher emend. Melkonian et Preisig		
COMMENT	Protist 152 (2), 139-147 (2001)		
FEATURES	21428165		
	Contact: Becker B		
	Botanisches Institut		
	Universitaet zu Koeln		
	Gyrfhofstr. 15, 50931 Koeln, Germany.		
	Location/Qualifiers		

source	1. 977		
	/organism="Scherffellia dubia"		
	/strain="SAG 40.89"		
	/db_xref="taxon:3190"		
	/clone="418"		
	/clone_lib="Scherffellia dubia pSPORT1"		
BASE COUNT	216 a	330 c	276 g 154 t
ORIGIN			
Query Match	22.6%;	Score 330.2;	DB 9; Length 977;
Best Local Similarity	64.5%;	Pred. No. 2.1e-76;	
Matches	522;	Conservative 1;	Mismatches 284; Indels 2; Gaps 2;
QY	481	ttcgaacacagccggtgctgttccagagccacacagaagctgacaaacctcgaatccctgc	540
Db	93	ttcccccaggagggcaccatccccaaacccggagtcaccgacaaacccagagttcaagatc	152
QY	541	gtttctgctacacacacagcaggtcttccaaacacagaagaacacacacacacagttgc	600
Db	153	gtgctgacccctcatcaaggaggccctgcagaaaggacccccagaaagtaccacaaagatgcc	212
QY	601	gccgcatgaagcgtgttccgaagagacacaaacacaggtgtccaccgctcttaccagctc	660
Db	213	aagcccttgtaggctgtccagagagaccacacacccgggtgaagc-stttgtacgagatg	271
QY	661	gagaaggaggccaaactctctccagccatcaacgtcaacacacacacacacacacacac	720
Db	272	caggacacacgctgctgtgttcccccgcacacacacacacacacacacacacacacac	331
QY	721	aagttcgataacatagcgtgcgcgcacacacacacacacacacacacacacacacacac	780
Db	332	aagttcgataacatagcgtgcgcgcacacacacacacacacacacacacacacacacac	391
QY	781	gatgcatgacggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg	840
Db	392	gacgtcattgctgcagggcagacacacacacacacacacacacacacacacacacacac	451
QY	841	tgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	900
Db	452	tgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	511
QY	901	tgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	960
Db	512	tgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	571
QY	961	gatgcatgacggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg	1020
Db	572	tgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	631
QY	1021	gacgtcattgctgcagggcagacacacacacacacacacacacacacacacacacacac	1080
Db	632	gacgtcattgctgcagggcagacacacacacacacacacacacacacacacacacacac	691
QY	1081	acagatggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1140
Db	692	atggcggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	751
QY	1141	atgtgggaattccagatggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1200
Db	752	atgtgggaattccagatggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	810
QY	1201	cgtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1260
Db	811	cgtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	870
QY	1261	ctgctcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	1289
Db	871	atggcggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	899
RESULT	8		
BI244550			



library HVCDNA0004 (Blumeria challenged)  
/tissue\_type="seedling green leaf"  
/lab\_host="TJC121"  
/note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI;  
C.I. 16155 (Mla13) plants were greenhouse grown in the R  
Wise lab at Iowa State University, Ames, IA; 7 day old  
green seedlings were challenged with isolate A27 (AvrMla13  
) of Blumeria graminis f. sp. hordei, and leaves were  
harvested 20 and 24 hr post-inoculation and snap frozen;  
uninoculated leaves were harvested 20 hr post-inoculation  
(Wei, Wise). In the TJ Close lab at the University of  
California, Riverside, total RNA was prepared from each  
sample pool, equal quantities of all three RNA pools were  
combined, poly(A) RNA was purified from the mixture, one  
cDNA library was made, and 1 million pfu were in vivo  
excised to give phagescript SK(-) cDNA phagemids (Choi,  
Close). Phagemids were plated and picked at the Clemson  
University Genomics Institute (CUGI) (Begum, Palmer,  
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA  
sequencing and sequence analysis were performed at CUGI  
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).  
The sequence has been trimmed to remove vector sequence  
and contains a minimum of 100 bases of phred value 20 or  
above. For more details on library preparation and  
sequence analysis see  
http://www.genome.clemson.edu/projects/barley. To order  
this clone see http://www.genome.clemson.edu/orders Also  
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)  
205 a 276 c 234 g 193 t

Query Match 21.6%; Score 316.2; DB 10; Length 908;  
Best Local Similarity 65.9%; Pred. No. 1e-72;  
Matches 475; Conservative 0; Mismatches 243; Indels 3; Gaps 1;

QY 540 cgtcttctacacacagcaggtcttccaaagacaaagacacagtgctgcacacagtgctgc 599  
DB 9 CGTCTCACCATCATCCGGCAGCGCTCAAGACCGACGCGGCTCAAGAGGCTCTACACAGT 68  
QY 600 tgcgggcatgaacggtgtttccgaagagacaaacacaggtgtccacgcgtctaccagct 659  
DB 69 GGAGAGGCTCGTCGGTGTCTCCGAGGAGACACACAGCGGCTCAAGAGGCTCTACACAGT 128  
QY 660 cgagaaggaggcaaacctctctccagccatcaacgtcaacgacgctgttcaaaagtc 719  
DB 129 GCAGAGGTCGGCAGCCCTCTCTTCCCGCCCATCAAGCTCAACGACGCTCCGTCACCAAGAG 188  
QY 720 caagtcgataacatctacggtgcgcgcacactcccttatcgatgggtatcaacgctgttc 779  
DB 189 CAAGTTTGACACACCTTACGGTTGCCGTCTCCTCTCCTCTCCTCTCCTCAGGCTCTCAGGCGCAC 248  
QY 780 cgtatgtatgatcgccggaagacagctctcgtcgttggttaagcgatgtcgcaagg 839  
DB 249 TGATGTATGATGATGCGGTAAGGTCGCCGCTGGTCTCGGTTATGTTGTTGCAAGGG 308  
QY 840 ctgcgctcaatccctccgtggccaaagcgctcgtcttatcatcacagaactccgacccaat 899  
DB 309 CTGTGCGCGCGGCACTCAACGAGCTGGTGGCCGTGTGATCGTGCACAGAGATCGACCCCAT 368  
QY 900 ctgcgctctccagggtgccatggaaggtaccagggtccgcgcgcgcgcgcgcgcgcgcgc 959  
DB 369 CTGTGCTCTCAGGCGCTGATGGAGGCACTCCAGATCTCTCACCCTGGAGGATGTGTCTC 428  
QY 960 ggtatgtatgtatgttgcacagcagaagaaactcgatctctctgttgatgat 1019  
DB 429 TGAGGCTGATATCTTTGTGACCAACACCGGAAACAGGACATCATATGTTGGACCAT 488  
QY 1020 gggccagatgaaggatgctattgtcgttaacatcgccacttcgataacgaattga 1079

Db 489 GAGGAAGATGAAGAACAAACGCGCATTTGTCGCAACATTTGGTCTACTTTGACAAACGAGATCGA 548  
QY 1080 tacagatggcctcatgaataccacaggcatcaagcacatcccaatcaagccagaatacga 1139  
Db 549 CATGAACGGCTTTGAGACCTACCCGGTGTCAAGCGCATCAACCATCAAGCCCAACTGA 608  
QY 1140 catgtgggaattccacagatggcca---cgctatccctctctctgtgtagggcgccctct 1196  
Db 609 CGCTGGGTCTTCCCTGAACCAAGACCGCATCATTTGTTCTTCTGCTGAGGTCGCTCTGAT 668  
QY 1197 taaccttgctgcgtacaggtcacccatcttctcgttatgtcaatgctcattcacaaacca 1256  
Db 669 GAACCTTGGATGTGCACGTGCCACCCGCCACCTTTGTCATGTCCTGCTCTTCTCACTAACCA 728  
QY 1257 g 1257  
Db 729 G 729

RESULT 10  
BF253807  
LOCUS  
DEFINITION  
HVSMEF0002C21f Hordeum vulgare seedling root EST library HVCDNA0007  
(Etiolated and unstressed) Hordeum vulgare cDNA clone  
HVSMEF0002C21f, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BF253807  
BF253807.2 GI:13116738  
EST.  
barley.  
Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Hordeum.  
REFERENCE  
AUTHORS  
1 (bases 1 to 867)  
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu  
Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton  
R.D., Oates, R. and Main, D.  
Development of a genetically and physically anchored EST resource  
for barley genomics: More unstressed seedling root cDNA library  
Unpublished (2001)  
On Nov 16, 2000 this sequence version replaced gi:11182912.  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: AATTAAACCTCACTAAAGGG  
High quality sequence stop: 730.  
Location/Qualifiers  
1. .867  
/organism="Hordeum vulgare"  
/cultivar="Morex"  
/db\_xref="taxon:4513"  
/clone="HVSMEF0002C21f"  
/clone\_lib="Hordeum vulgare seedling root EST library  
HVCDNA0007 (Etiolated and unstressed)"  
/tissue\_type="Seedling root"  
/lab\_host="TJC121"  
/note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI;  
Seeds were surface sterilized then germinated under axenic  
conditions in the dark at room temperature on filter paper  
with water, nystatin and cefotaxime in covered  
crystallization dishes. Five-day old seedling roots were  
then harvested, total RNA was prepared, poly(A) RNA was  
purified, one primary unamplified cDNA library was made,  
and 1 million pfu were in vivo excised to give phagescript  
SK(-) cDNA phagemids. These steps were performed in the TJ  
Close laboratory at the University of California,  
Riverside (Choi, Close, Fenton). Phagemids were plated and  
picked at the Clemson University Genomics Institute (CUGI)





```

Db 481 ACCACGGTAACAAGGACATCATCATGTTGACCAATGAGGAAGATGAAGAACAATGCC 540
QY 1042 attgtcgttaacatcgccacttcgataacgaattgatacagatggcctcatgaataac 1101
Db 541 ATTGTCGCAACATTTGGTTCACCTTGACAAAGAAATCGACATGCATGGTCTTGAACCTTC 600
QY 1102 ccaggcaataagacacatcccccaatacaagccagaatacagacatgtgggaattcccagatggc 1161
Db 601 CCRGGTGTGAAGAGGATGACAAATCAAGCCACAAACCGACAGATGGGTCTCCGAGACACC 660
QY 1162 cac---gctatccctctcttgcgtgagggccgcctcttaaccttggctgagctcaggt 1218
Db 661 AACAGTGGCATCATGCTGTGGCCGAGGGTGTCTCATGAACCTTGGGATGTGCCACTGGA 720
QY 1219 caccatcttctgattgcaatgtcatgtcatcaacaaacagacactgcgtcagctcgacctc 1278
Db 721 CACCCCAAGTTTTGTGATGCTCTTCTCTTCTACCAAGTCATTTGCCCAACTCAAGTTG 780
QY 1279 t 1279
Db 781 T 781

RESULT 13
BG444868
LOCUS GA_Ea0025010f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION arboreum cDNA clone GA_Ea0025010f, mRNA sequence.
ACCESSION BG444868
VERSION BG444868.1 GI:13354520
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE 1 (bases 1 to 916)
AUTHORS Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE An integrated analysis of the genetics, development, and evolution
of the cotton fiber
JOURNAL unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 639.
Location/Qualifiers
FEATURES
source
1..916
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0025010f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 225 a 209 c 264 g 218 t
ORIGIN

Query Match 20.8%; Score 303.4; DB 10; Length 916;
Best Local Similarity 63.0%; Pred. No. 2.5e-69;
Matches 521; Conservative 0; Mismatches 296; Indels 10; Gaps 3;
QY 625 gagacaacaaggtgtccaccgctctaccagctcagagaggggcaaacctctcttc 684

```

```

Db 12 GAAATACAAACGGTGTAAAGAGCTTTATCAGATGAGGCCAATGGAACCTTGTGTTC 71
QY 685 ccagccatcaacgctcaacgacgctgttcaaaagtcctcaagttcgaatacatcaggtgc 744
Db 72 CCTGTATTAATGTCAATGACTCTGTCAACCAAGAGCAAGTTCGATTAACCTTGTATGGATGC 131
QY 745 ggcacatccctctatgattggtatcaaccgtgtcccgatgtcatgtatgctggcggaagaca 804
Db 132 CBTCACTCTCTTCCCGATGGTTGATGAGAGCTACCCATGTATGATGCTGCTGGTAAAGTTC 191
QY 805 gctctcgtcatgggttacggcgatgctggcaagggtgcgctcaatccctccgtgggcaaa 864
Db 192 GCTGTGTCTGTTGTTATGTTGATGTCGCAAGGGTGTGCTGCTGCTTGAAGCAAGCC 251
QY 865 ggcgtcgcgttatcatcatcacagaactcgaccgaatctgcgtctccaggtgcatgaa 924
Db 252 GGTGCTCCGCTCATTTGTCAACGAGATTGATCCCATGTGCCCTTCAGGCTCTCATGAA 311
QY 925 ggtaccagggtccgcgcatcgaggaagtcgcaagatgctcgaatgtcgatatcttgcgttacatgc 984
Db 312 GGACTTCAAGTTTTGACCTTGAGGATGTGTCGAGGCTGATATCTTTGTCCACCAGC 371
QY 985 acaggaactgcgtatcatctctgttgacatgatggccagatgaaggataaaggctatt 1044
Db 372 ACTGTTAAACAAGGACATCATCATGTTTAAACCATGAGGAAGATGAAGAACAATGCCATT 431
QY 1045 gtcggttaacatcgccacttcgataacgaattgatacagatggcctcatgaataccca 1104
Db 432 GTTTGCAACATCGGTTCACCTTCGACAAATGAATTGACATGCTCGCTCTCGAGAATTAATCCT 491
QY 1105 ggcatacagcatccatcccaatcaagccagaatacagacatgtgggaattcccagatggcca - 1163
Db 492 GGTGCTCAAGCGCATTACCATCAAGCTCAACCGGATAGGTGGTCTTCCCTGAACCAAC 551
QY 1164 --cgctatcctctcttgcgtgagggccgctctttaaacttggctgctgctacaggtcac 1221
Db 552 ACCGGCATCGTGTGTGTTGGCTGAAGAGCTCTCATGAACCTTGGTGTGCTACTGGCCAC 611
QY 1222 ccattcttcgttatgtaaatgtcattcacaacagacactgcgtcagctcgacctetac 1281
Db 612 CCTAGCTTTGGCATGTCTATGCTGTTCACTAACCCAGGTGATCGCCAGCTGGAGCTGTGG 671
QY 1282 gaaag-----agagaaatctcgagaagaaggtttacacacttcggaagcatctcgat 1335
Db 672 TAAGAGAATGCAACAGCAAGTGCAGAGAAGAAAGAGGGGGTTCCTTAACACCTTGAT 731
QY 1336 gaagaagtgcgtgcctccacctcggatctctcgatgtccaccttacaagcttacacag 1395
Db 732 GAGAAAGTGGTGTCTTTAGCTTTGAAAGCT-GAGGGGTACCTACCAAGCTCACCAGA 790
QY 1396 aagcaggctgaactacatcaacgcttccagttgaggggtccttacaagtc 1442
Db 791 GATCAAAAGCAGCTAAATTAAGCGGGGCCCTAGGGGAGCTTAAAGGCC 837

RESULT 14
BF863943
LOCUS BF863943
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BF863943
VERSION BF863943.1 GI:12254087
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 769)
AUTHORS Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C.,
Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
ANALYSES OF THE CHLAMYDOMONAS REINHARDTII GENOME: A MODEL,
UNICELLULAR SYSTEM FOR ANALYZING GENE FUNCTION AND REGULATION IN

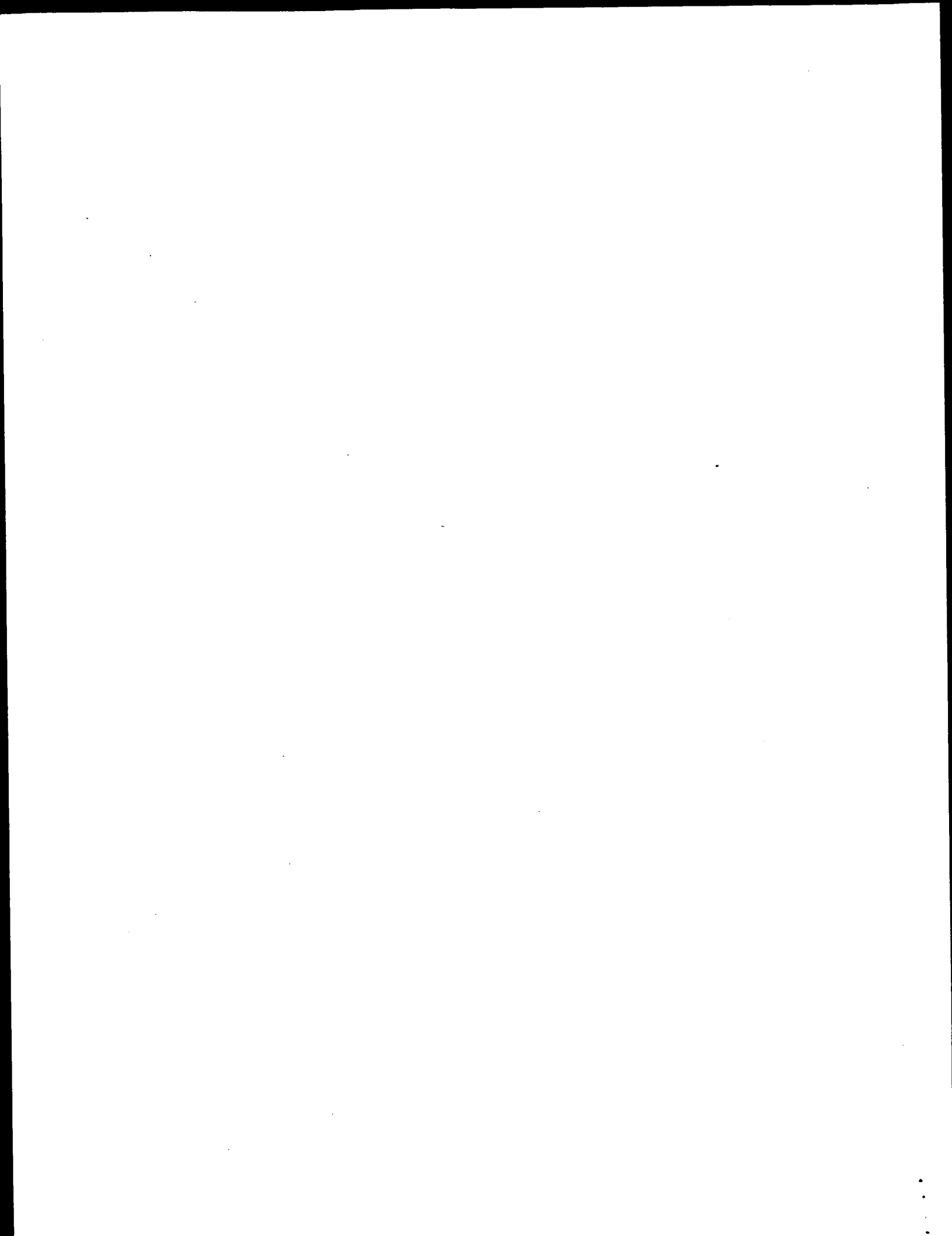
```











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:44:41 ; Search time 268.29 Seconds  
(without alignments)  
9349.629 Million cell updates/sec

Title: US-09-759-990-1

Perfect score: 1461

Sequence: 1 atggttgcaaatcacctac.....ctgatgcttaccgtatttaa 1461

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Capext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1459.4	99.9	1461	22	Trichomonas vagina
2	1440.2	98.6	1461	22	Trichomonas vagina
3	596.4	40.8	1557	22	Corynebacterium gl
4	596.4	40.8	1557	23	C. glutamicum gene
5	596.4	40.8	349980	22	C glutamicum codin
6	594	40.7	1422	22	C glutamicum codin
7	535.4	36.6	1488	21	M. tuberculosis an
8	529.6	36.2	1396	22	Corynebacterium gl
9	529.6	36.2	1396	23	C. glutamicum gene

10	508.8	34.8	1865	21	AAC44037	Zea mays DNA fragm
11	506.6	34.7	1727	21	AAC46315	Arabidopsis thalia
12	497.8	34.1	1728	21	AAC39836	Arabidopsis thalia
13	489	33.5	1767	17	AAT44513	Asparagus S-adenos
14	488.4	33.4	1812	13	AAQ29419	Nicotiana tabacum
15	488.4	33.4	1812	17	AAT23194	S-adenosylhomocyst
16	471.4	32.3	1494	21	AAC46817	Arabidopsis thalia
17	375.4	25.7	4411529	22	AAI99682	Mycobacterium tube
C	331	22.7	2200	21	AAC77823	Human cancer assoc
C	302	20.7	513445	22	AAI61373	Soybean 318013 reg
20	297.8	20.4	1941	17	AAT44515	Arabidopsis S-aden
21	293.4	20.1	1398	21	AAT98330	A. thaliana gene i
C	288.4	19.7	720	22	AAH65803	C glutamicum codin
C	278.6	19.1	708	22	AAH65802	C glutamicum codin
24	244	16.7	2563	19	AAV28617	Nucleotide sequenc
25	244	16.7	3616	20	AAV73924	Human SAHH DNA #1.
26	233.6	16.0	1566	23	ABL03693	Drosophila melanog
27	223.2	15.3	2211	22	AAC91228	Human S-adenosylho
C	179.2	12.3	5030	23	AAV76216	DNA encoding novel
C	161.2	11.0	4406	23	ABL03692	Drosophila melanog
C	149.8	10.3	58909	22	AAF28543	Genomic fragment #
C	146.8	10.0	349980	22	AAH41226	Pyrococcus abyssi
32	137	9.4	793	19	AAV28651	Ripening banana pu
C	125.2	8.6	1949	21	AAV07609	Fusarium venenatum
34	121	8.3	363	21	AAC01924	Human secreted pro
35	120.8	8.3	289	16	AAQ98682	Human SAHH cDNA.
36	120.8	8.3	656	21	AAF14585	Aspergillus oryzae
37	108.2	7.4	371	21	AAA31988	Plant microsateili
38	108.2	7.4	722	19	AAV28650	Ripening banana pu
39	106.8	7.3	374	21	AAA31802	Plant microsateili
40	105	7.2	285	16	AAQ98681	Plant microsateili
41	99	6.8	3795	22	AAC87646	SAHH-GSE. Mus sp.
42	96	6.6	352	21	AAA31634	Brassica napus myo
C	95.4	6.5	573	22	AAI90254	Plant microsateili
C	91.4	6.3	1000	22	AAF91420	Human polynucleoti
45	87	6.0	4761	23	ABL14956	Moraxella catarrha
						Drosophila melanog

#### ALIGNMENTS

RESULT 1

ID AAS09278 standard; DNA; 1461 BP.

AC AAS09278;

XX

XX

DT 24-OCT-2001 (first entry)

DE Trichomonas vaginalis recombinant His-SAHH gene sequence.

XX

KW SAHH; S-adenosyl homocysteine; S-adenosyl homocysteine hydrolase;

KW S-adenosylmethionine; SAM; glycine N-methyltransferase; GMT; His-SAHH;

KW cancer; malaria; arthritis; SAH; mutant; ds.

XX

OS Trichomonas vaginalis.

OS Synthetic.

XX

FT Key

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

FT mutation

Location/Qualifiers  
replace (19,G)  
/\*tag= a  
replace (201,G)  
/\*tag= b  
replace (207,T)  
/\*tag= c  
replace (210,T)  
/\*tag= d  
replace (501,C)  
/\*tag= e  
replace (744,T)  
/\*tag= f  
replace (834,G)  
/\*tag= g

mutation replace (897,T)  
 /\*tag= h  
 mutation replace (917,T)  
 /\*tag= i  
 mutation replace (1206,T)  
 /\*tag= j  
 mutation replace (1304,T)  
 /\*tag= k  
 mutation replace (1346,T)  
 /\*tag= l

WO200151651-A2.

19-JUL-2001.

12-JAN-2001; 2001WO-US01114.

14-JAN-2000; 2000US-0176444.

(ANTI-) ANTICANCER INC.

Hoffman RM, Xu M, Han Q;

WPI; 2001-451863/48.

Assessing therapeutic levels of S-adenosylmethionine comprises

measuring reaction products in sample containing glycine

N-methyltransferase, (His) S-adenosyl homocysteine hydrolase and

glycine

Claim 6; Fig 6A-6B; 28pp; English.

The present sequence for His-SAHH represents a novel recombinant gene  
 sequence for Trichomonas vaginalis S-adenosyl homocysteine  
 (SAHH, S-adenosyl homocysteine hydrolase). The wild type SAHH gene  
 sequence (AAS09277) is used to generate the His-SAHH gene sequence  
 which comprises various point mutations and an additional polynucleotide  
 sequence which encodes for an extra 6 N-terminal histidine residues.

The invention relates to a method for assessing therapeutic

levels of S-adenosylmethionine (SAM) in a biological fluid sample

comprising measuring one or more reaction products in a sample

containing glycine N-methyltransferase (GNF), SAHH or His-SAHH, and

glycine, where the level of one or more products is directly

proportional to the level of SAM in the sample. The method is useful

for assaying therapeutic levels of SAM in a biological sample and may

be used as a part of a diagnostic protocol or as part of a therapeutic

protocol, where conditions or progress of the therapy may be monitored.

SAHH or His-SAHH may be used as a reagent, particularly in screening for

inhibitors and inactivators of the enzyme, for use as reagents themselves

as potential therapeutics, e.g. in cancer, malaria, arthritis and other

diseases. Recombinant SAHH may be used as a therapeutic cancer gene in

combination with S-adenosyl homocysteine (SAH) analogues.

XX Sequence 1461 BP; 372 A; 427 C; 335 G; 327 T; 0 other;

Query Match 99.9%; Score 1459.4; DB 22; Length 1461;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggcttgcaaatcacctactagggtgctcattcgtagtacagaaatgccacatcaacctc 60

Db 1 atggcttgcaaatcacctactagggtgctcattcgtagtacagaaatgccacatcaacctc 60

QY 61 catgtctccggcgttaaggaaactacccttctgtgagaagaaatgccaggtcttatggtt 120

Db 61 catgtctccggcgttaaggaaactacccttctgtgagaagaaatgccaggtcttatggtt 120

QY 121 ctctgtgagcgttatctccgtctcttaagcattgaagggtgtcagaatctctgttccctc 180

Db 121 ctctgtgagcgttatctccgtctcttaagcattgaagggtgtcagaatctctgttccctc 180

QY 181 cacatgcagctccagagacgcgtctctatcagacactcacagctcttggctgatgctc 240

Db 181 cacatgcagctccagagacgcgtctctatcagacactcacagctcttggctgatgctc 240

QY 241 agatgggcttctcgaacatctctctacacagatacagccgctgctgtctatctgttgc 300

Db 241 agatgggcttctcgaacatctctctacacagatacagccgctgctgtctatctgttgc 300

QY 301 ggcccaacagcgcacaccagagaagccagcgggtatccccagttctccctggaaggcgaa 360

Db 301 ggcccaacagcgcacaccagagaagccagcgggtatccccagttctccctggaaggcgaa 360

QY 361 acatcccaagaatactgggagaacacataccgcgtctctcacatggccagatggtcaaggc 420

Db 361 acatcccaagaatactgggagaacacataccgcgtctctcacatggccagatggtcaaggc 420

QY 421 ccacagcaggttgcgatgagtggtgatgtctacactctctctccaaagggtctcgaa 480

Db 421 ccacagcaggttgcgatgagtggtgatgtctacactctctctccaaagggtctcgaa 480

QY 481 ttgaaacagccggtgctgttccagagcccaacagagctgacaaacctcgaaacctgcgc 540

Db 481 ttgaaacagccggtgctgttccagagcccaacagagctgacaaacctcgaaacctgcgc 540

QY 541 gttctgtctactcaagcaggtcttccaaacagacagaacacacacagctggcacagttgct 600

Db 541 gttctgtctactcaagcaggtcttccaaacagacagaacacacacagctggcacagttgct 600

QY 601 gccggcatgaacggtgtttccgaagagacaaacacaggtgtccaccgcctctaccagctc 660

Db 601 gccggcatgaacggtgtttccgaagagacaaacacaggtgtccaccgcctctaccagctc 660

QY 661 gagaagggggcaaaactctctccagccatacaacgctcaaacgacgctgttacaagtcc 720

Db 661 gagaagggggcaaaactctctccagccatacaacgctcaaacgacgctgttacaagtcc 720

QY 721 aagttcgataacatagcgtccgcgcactccctctatcgatggtatcaaccgctctcc 780

Db 721 aagttcgataacatagcgtccgcgcactccctctatcgatggtatcaaccgctctcc 780

QY 781 gatgtcatgatcggcggcaagacagctctcgtcatgggttacggcgagtgctggcaaggc 840

Db 781 gatgtcatgatcggcggcaagacagctctcgtcatgggttacggcgagtgctggcaaggc 840

QY 841 tgcgctcaatccctccgtggcgaaggcgtcgcgttatctatcatcacagaactcgaccatc 900

Db 841 tgcgctcaatccctccgtggcgaaggcgtcgcgttatctatcatcacagaactcgaccatc 900

QY 901 tgcgctccagcgtccatgaaggctaccaggtccgcgcctcgcagatcgaggagtgctcaag 960

Db 901 tgcgctccagcgtccatgaaggctaccaggtccgcgcctcgcagatcgaggagtgctcaag 960

QY 961 gatgtcgatatcttctgtacatgcacaggaactcgatcatctctctgttgacatgatg 1020

Db 961 gatgtcgatatcttctgtacatgcacaggaactcgatcatctctctgttgacatgatg 1020

QY 1021 gccacagatgaagataaaggctattgtcgttaacatcgccacttcgataacgaattgat 1080

Db 1021 gccacagatgaagataaaggctattgtcgttaacatcgccacttcgataacgaattgat 1080

QY 1081 acagatggcctcatgaataccacagcgcatacgaacatcccaatcaagcagatacagac 1140

Db 1081 acagatggcctcatgaataccacagcgcatacgaacatcccaatcaagcagatacagac 1140

QY 1141 atgtgggaattcccagatggccacgctatctctctcttctgtgagggccgcctcttaac 1200

Db 1141 atgtgggaattcccagatggccacgctatctctctcttctgtgagggccgcctcttaac 1200

QY 1201 ctgtgctgcgtacaggttaacccatcttctgttatgtcaatgtcaattcaaacacagaca 1260

Db 1201 ctgtgctgcgtacaggttaacccatcttctgttatgtcaatgtcaattcaaacacagaca 1260

QY 1261 ctgcgtcagctcagcctctacgaaagagaggaaatctcgagaagaagggtttacacatt 1320

Db 1261 ctgcctcagctcgcagcctctcgaagaaagaggaatctcgaagaaggtttacacact 1320  
QY 1321 ccgaagatctcgtatgaagaagtcgctccctccacctcgcgtctcgtatccacact 1380  
Db 1321 ccgaagatctcgtatgaagaagtcgctccctccacctcgcgtctcgtatccacact 1380  
QY 1381 acaagcttccacgaagaagcaggtctacatcaacgttccagttcaggtccttacaag 1440  
Db 1381 acaagcttccacgaagaagcaggtctacatcaacgttccagttcaggtccttacaag 1440  
QY 1441 tctgatctacagttatata 1461  
Db 1441 tctgatctacagttatata 1461

RESULT 2  
AAS09277  
ID AAS09277 standard; DNA; 1461 BP.  
AC AAS09277;  
XX  
XX 24-OCT-2001 (first entry)  
DE Trichomonas vaginalis wild type SAHH gene sequence.  
XX  
XX SAHH; S-adenosyl homocysteine; S-adenosyl homocysteine hydrolase;  
KW S-adenosylmethionine; SAM; glycine N-methyltransferase; GMT; His-SAHH;  
KW cancer; malaria; arthritis; SAH; ds.  
XX  
XX Trichomonas vaginalis.  
OS  
FH Key Location/Qualifiers  
FH mutation replace (19,A)  
FT /\*tag- a  
FT mutation replace (201,C)  
FT /\*tag- b  
FT mutation replace (207,C)  
FT /\*tag- c  
FT mutation replace (210,C)  
FT /\*tag- d  
FT mutation replace (501,T)  
FT /\*tag- e  
FT mutation replace (744,C)  
FT /\*tag- f  
FT mutation replace (834,C)  
FT /\*tag- g  
FT mutation replace (897,A)  
FT /\*tag- h  
FT mutation replace (917,C)  
FT /\*tag- i  
FT mutation replace (1206,C)  
FT /\*tag- j  
FT mutation replace (1304,A)  
FT /\*tag- k  
FT mutation replace (1346,C)  
FT /\*tag- l  
XX WO200151651-A2.  
XX PN 19-JUL-2001.  
XX PD 12-JAN-2001; 2001WO-US01114.  
XX PF 14-JAN-2000; 2000US-0176444.  
XX PR  
XX (ANTI-) ANTICANCER INC.  
XX PI Hoffman RM, Xu M, Han Q;  
XX WPI; 2001-451863/48.  
XX  
XX Assessing therapeutic levels of S-adenosylmethionine comprises  
PT measuring reaction products in sample containing glycine

PT N-methyltransferase, (His) S-adenosyl homocysteine hydrolase and  
PT glycine  
PS Disclosure; Fig 6A-6B; 28pp; English.  
XX

The present sequence represents the gene sequence for Trichomonas vaginalis wild type S-adenosyl homocysteine hydrolase (SAHH, S-adenosyl homocysteine hydrolase). The wild type SAHH gene sequence is used to generate a novel recombinant SAHH gene sequence, His-SAHH, (AAS09278) which comprises various point mutations and an additional polynucleotide sequence which encodes for an extra 6 N-terminal histidine residues. The invention relates to a method for assessing therapeutic levels of S-adenosylmethionine (SAM) in a biological fluid sample comprising measuring one or more reaction products in a sample containing glycine N-methyltransferase (GMT), SAHH or His-SAHH, and glycine, where the level of one or more products is directly proportional to the level of SAM in the sample. The method is useful for assaying therapeutic levels of SAM in a biological sample and may be used as a part of a diagnostic protocol or as part of a therapeutic protocol, where conditions or progress of the therapy may be monitored. SAHH or His-SAHH may be used as a reagent, particularly in screening for inhibitors and activators of the enzyme for use as reagents themselves as potential therapeutic agents, e.g. in cancer, malaria, arthritis and other diseases. Recombinant SAHH may be used as a therapeutic cancer gene in combination with S-adenosyl homocysteine (SAH) analogues.

Sequence 1461 BP; 369 A; 420 C; 338 G; 334 T; 0 other;

Query Match 98.6%; Score 1440.2; DB 22; Length 1461;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1448; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 atggttgcaatccactacggtggtcctcagatcagaaatgcagacataactc 60  
Db 1 atggttgcaatccactacggtggtcctcagatcagaaatgcagacataactc 60  
QY 61 catgttcgcgcgtaagaacttacccttgctgaaagaatgcaggtctatagtt 120  
Db 61 catgttcgcgcgtaagaacttacccttgctgaaagaatgcaggtctatagtt 120  
QY 121 ctctcgtacggtatctccgctcctaagcattgaagggtgcagaatctctgtcc 180  
Db 121 ctctcgtacggtatctccgctcctaagcattgaagggtgcagaatctctgtcc 180  
QY 181 cacatgacgtccagaaagccgctcctatcgagaactcaagctctgtgctgattc 240  
Db 181 cacatgacgtccagaaagccgctcctatcgagaactcaagctctgtgctgattc 240  
QY 241 agatggcttcctcgaacatctctctacacagaatagacgcgtgctctgttc 300  
Db 241 agatggcttcctcgaacatctctctacacagaatagacgcgtgctctgttc 300  
QY 301 ggcaccaagcagacacagagaagccagccggtatccagttcgcctggaagcgaa 360  
Db 301 ggcaccaagcagacacagagaagccagccggtatccagttcgcctggaagcgaa 360  
QY 361 acactccagaatactggaagaacataccggtctccatctcagatgcagatgtaaggc 420  
Db 361 acactccagaatactggaagaacataccggtctccatctcagatgcagatgtaaggc 420  
QY 421 ccacagcaggttgcgtatggtggtggtggtggtggtggtggtggtggtggtggt 480  
Db 421 ccacagcaggttgcgtatggtggtggtggtggtggtggtggtggtggtggtggt 480  
QY 481 ttcgaagaagcgggtggttccagagccaaagaagctgcacactcgaatccgctgac 540  
Db 481 ttcgaagaagcgggtggttccagagccaaagaagctgcacactcgaatccgctgac 540  
QY 541 gtcttgcataactcaagcaggtctcacaacagaagaacacactgacacacttctc 600  
Db 541 gtcttgcataactcaagcaggtctcacaacagaagaacacactgacacacttctc 600

QY 601 gccggcgtgacgggtgttccgaagagacaacaggtgtccaccgctctaccagctc 660  
 |||||  
 Db 601 gccggcgtgacgggtgttccgaagagacaacaggtgtccaccgctctaccagctc 660  
 QY 661 gagaagagagggcaactctcttcccaagcatcaacgtcaacagcgctgttaacagtc 720  
 |||||  
 Db 661 gagaagagagggcaactctcttcccaagcatcaacgtcaacagcgctgttaacagtc 720  
 QY 721 aagtcgataacatctacagcgctgcgcacccctctatcgatgtgtatcaacgctgtcttc 780  
 |||||  
 Db 721 aagtcgataacatctacagcgctgcgcacccctctatcgatgtgtatcaacgctgtcttc 780  
 QY 781 gatgtcatgtatcgccggaagacaacgtctctgtcatgtgttaagcgatgtccgaagggc 840  
 |||||  
 Db 781 gatgtcatgtatcgccggaagacaacgtctctgtcatgtgttaagcgatgtccgaagggc 840  
 QY 841 tgcgctcaatccctccggtggccaagcgctcgttatcatcatcaacaactcgaccatc 900  
 |||||  
 Db 841 tgcgctcaatccctccggtggccaagcgctcgttatcatcatcaacaactcgaccatc 900  
 QY 901 tgcgctcccaagcgctgtcgaagagcaacaggtccgcgcacgtcgaagagtcgttaag 960  
 |||||  
 Db 901 tgcgctcccaagcgctgtcgaagagcaacaggtccgcgcacgtcgaagagtcgttaag 960  
 QY 961 gatgtcatgtatctgttcatcatgtcacaaggaactcgatcatctctgttgatcatgtg 1020  
 |||||  
 Db 961 gatgtcatgtatctgttcatcatgtcacaaggaactcgatcatctctgttgatcatgtg 1020  
 QY 1021 gcccgagatgaagagataagagctatgttcgttaacatcgccactcgataacgaatgtgat 1080  
 |||||  
 Db 1021 gcccgagatgaagagataagagctatgttcgttaacatcgccactcgataacgaatgtgat 1080  
 QY 1081 acagatgtgcctcatatgaataacccaagcatcaagacaatcccaatcaagcgagaatacagc 1140  
 |||||  
 Db 1081 acagatgtgcctcatatgaataacccaagcatcaagacaatcccaatcaagcgagaatacagc 1140  
 QY 1141 atgttggaatcccaagatgtgcacagctatccctctctgtcgaagggcgccctcttaac 1200  
 |||||  
 Db 1141 atgttggaatcccaagatgtgcacagctatccctctctgtcgaagggcgccctcttaac 1200  
 QY 1201 ctgtgctgcgtcaagagtgtaaccatcttctgttatgtcatgtcatcaacaacagaca 1260  
 |||||  
 Db 1201 ctgtgctgcgtcaagagtgtaaccatcttctgttatgtcatgtcatcaacaacagaca 1260  
 QY 1261 ctgcgtcagctcagctctcaagaaagaggaatcctcgagaagaggtttacacactt 1320  
 |||||  
 Db 1261 ctgcgtcagctcagctctcaagaaagaggaatcctcgagaagaggtttacacactt 1320  
 QY 1321 ccgaagcatctcgaatgaagaagtcgctcgcctccactcgatctctcgtatgtccactt 1380  
 |||||  
 Db 1321 ccgaagcatctcgaatgaagaagtcgctcgcctccactcgatctctcgtatgtccactt 1380  
 QY 1381 acaaaagcttaacagaagagcgctgactatacaacgttccagttgaggttcccttaag 1440  
 |||||  
 Db 1381 acaaaagcttaacagaagagcgctgactatacaacgttccagttgaggttcccttaag 1440  
 QY 1441 tctgagcttaccgttatata 1461  
 |||||  
 Db 1441 tctgagcttaccgttatata 1461  
 RESULT 3  
 AAF71862  
 ID AAF71862 standard; DNA; 1557 BP.  
 XX  
 AC AAF71862;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 DE Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:219.  
 XX  
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;  
 fine chemical production; microorganism; organic acid; nucleoside;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
 KW Lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
 KW carbohydrate; aromatic compound; cofactor; polypeptide; enzyme; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN WO200100843-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000MO-IB00923.  
 XX  
 XX 25-JUN-1999; 99US-0141031.  
 PR 01-JUL-1999; 99DE-1030476.  
 PR 02-JUL-1999; 99US-0142101.  
 PR 08-JUL-1999; 99DE-1031415.  
 PR 08-JUL-1999; 99DE-1031418.  
 PR 08-JUL-1999; 99DE-1031419.  
 PR 08-JUL-1999; 99DE-1031420.  
 PR 08-JUL-1999; 99DE-1031424.  
 PR 08-JUL-1999; 99DE-1031428.  
 PR 08-JUL-1999; 99DE-1031434.  
 PR 08-JUL-1999; 99DE-1031435.  
 PR 08-JUL-1999; 99DE-1031443.  
 PR 08-JUL-1999; 99DE-1031453.  
 PR 08-JUL-1999; 99DE-1031457.  
 PR 08-JUL-1999; 99DE-1031465.  
 PR 08-JUL-1999; 99DE-1031478.  
 PR 08-JUL-1999; 99DE-1031510.  
 PR 08-JUL-1999; 99DE-1031541.  
 PR 08-JUL-1999; 99DE-1031573.  
 PR 08-JUL-1999; 99DE-1031592.  
 PR 08-JUL-1999; 99DE-1031632.  
 PR 08-JUL-1999; 99DE-1031634.  
 PR 08-JUL-1999; 99DE-1031636.  
 PR 08-JUL-1999; 99DE-1032125.  
 PR 08-JUL-1999; 99DE-1032126.  
 PR 08-JUL-1999; 99DE-1032130.  
 PR 08-JUL-1999; 99DE-1032186.  
 PR 08-JUL-1999; 99DE-1032206.  
 PR 08-JUL-1999; 99DE-1032227.  
 PR 08-JUL-1999; 99DE-1032228.  
 PR 08-JUL-1999; 99DE-1032229.  
 PR 08-JUL-1999; 99DE-1032230.  
 PR 14-JUL-1999; 99DE-1032922.  
 PR 14-JUL-1999; 99DE-1032926.  
 PR 14-JUL-1999; 99DE-1032928.  
 PR 14-JUL-1999; 99DE-1033004.  
 PR 14-JUL-1999; 99DE-1033005.  
 PR 14-JUL-1999; 99DE-1033006.  
 PR 14-JUL-1999; 99US-0148613.  
 PR 14-JUL-1999; 99DE-1040764.  
 PR 27-AUG-1999; 99DE-1040765.  
 PR 27-AUG-1999; 99DE-1040766.  
 PR 27-AUG-1999; 99DE-1040832.  
 PR 31-AUG-1999; 99DE-1041378.  
 PR 31-AUG-1999; 99DE-1041379.  
 PR 31-AUG-1999; 99DE-1041380.  
 PR 31-AUG-1999; 99DE-1041384.  
 PR 31-AUG-1999; 99DE-1041396.  
 PR 31-AUG-1999; 99DE-1042076.  
 PR 03-SEP-1999; 99DE-1042077.  
 PR 03-SEP-1999; 99DE-1042079.  
 PR 03-SEP-1999; 99DE-1042086.  
 PR 03-SEP-1999; 99DE-1042087.  
 PR 03-SEP-1999; 99DE-1042088.  
 PR 03-SEP-1999; 99DE-1042095.  
 PR 03-SEP-1999; 99DE-1042124.  
 PR 03-SEP-1999; 99DE-1042129.  
 PR 09-MAR-2000; 2000US-0187970.  
 XX  
 PA (BADI ) BASF AG.  
 XX

PI Pompeius M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;  
 XX WPI: 2001-137957/14.  
 DR P-PSDB: AAB79743.  
 XX  
 PT Nucleic acids from *Corynebacterium glutamicum* encoding metabolic  
 PT pathway proteins, useful for producing fine chemicals in  
 PT microorganisms, including organic acids, nonproteinogenic amino acids,  
 PT and purine and pyrimidine bases -  
 XX  
 XX Claim 3: Page 473-475; 1737p; English.  
 XX  
 CC AAF71753 to AAF72330 encode the *Corynebacterium glutamicum* metabolic  
 CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum  
 CC MP nucleic acids are useful for the production of fine chemicals  
 CC in microorganisms, including organic acids, nonproteinogenic amino  
 CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,  
 CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic  
 CC compounds, vitamins, cofactors, polyketides and enzymes.  
 CC  
 XX Sequence 1557 BP: 348 A; 440 C; 431 G; 338 T; 0 other;

Query Match 40.8%; Score 596.4; DB 22; Length 1557;  
 Best Local Similarity 64.5%; Pred. No. 7.3e-166;  
 Matches 923; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 32 tcgaatcagaatgcccagacatccatcgtcttcgagcgaaggaactacccctg 91  
 DB 114 tggactcagaatgcccagatcttcacagagagcagagcagcagatcgcctg 173  
 QY 92 ctgagaggaatgcccagatcttcgtcttcgagcgaaggaactacccctg 151  
 DB 174 cagaagatgagatgcccagatcttcacagagagcagaggaactcgcagagcagcct 233  
 QY 152 tgaaggggtcagaatctcgtcttcacacatgacagcagcagcagcagcagc 211  
 DB 234 tgaaggggtcagaatctcgtcttcacacatgacagcagcagcagcagcagc 293  
 QY 212 agacatcagaatctcgtcttcacacatgacagcagcagcagcagcagcagc 271  
 DB 294 agacatcagaatctcgtcttcacacatgacagcagcagcagcagcagcagc 353  
 QY 272 aagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 331  
 DB 354 agatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 410  
 QY 332 gtaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 391  
 DB 411 gtaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 470  
 QY 392 ggcctcacaatgcccagatgagcagcagcagcagcagcagcagcagcagc 451  
 DB 471 agatcacaatgcccagatgagcagcagcagcagcagcagcagcagcagc 527  
 QY 452 ctacactcacaatgcccagatgagcagcagcagcagcagcagcagcagcagc 511  
 DB 528 ccacatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 587  
 QY 512 cagaagctcagaacccagatgagcagcagcagcagcagcagcagcagcagc 571  
 DB 588 aggcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 647  
 QY 572 aagacaagaacacagcagcagcagcagcagcagcagcagcagcagcagcagc 631  
 DB 648 cagaagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 707  
 QY 632 caacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 691  
 DB 708 caacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 767  
 QY 692 tcaacgtaacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 751

DB 768 tgaacgtcaacagcagcagcagcagcagcagcagcagcagcagcagcagc 827  
 QY 752 cccttatcagatgatatcaacgcagcagcagcagcagcagcagcagcagcagc 811  
 DB 828 ccccttatcagatgatatcaacgcagcagcagcagcagcagcagcagcagcagc 887  
 QY 812 tcaaggttcaacagcagcagcagcagcagcagcagcagcagcagcagcagc 871  
 DB 888 tccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 947  
 QY 872 ggccttatcacaacagcagcagcagcagcagcagcagcagcagcagcagcagc 931  
 DB 948 ggccttatcacaacagcagcagcagcagcagcagcagcagcagcagcagcagc 1007  
 QY 932 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 991  
 DB 1008 cgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1067  
 QY 992 actgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1051  
 DB 1068 acaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1127  
 QY 1052 acatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1111  
 DB 1128 acatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1187  
 QY 1112 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1171  
 DB 1188 cccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1247  
 QY 1172 tcccttatcagatgatatcaacgcagcagcagcagcagcagcagcagcagcagc 1231  
 DB 1248 tcccttatcagatgatatcaacgcagcagcagcagcagcagcagcagcagcagc 1307  
 QY 1232 ttatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1291  
 DB 1308 tcccttatcagatgatatcaacgcagcagcagcagcagcagcagcagcagcagc 1367  
 QY 1292 gaaatcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1351  
 DB 1368 gaaatcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1427  
 QY 1352 tccatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1411  
 DB 1428 tccatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1487  
 QY 1412 tcaacgtaacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1461  
 DB 1488 tcaacgtaacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1537

RESULT 4  
 AAS96121  
 ID AAS96121 standard; DNA; 1557 BP.  
 XX  
 AC AAS96121:  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE C. glutamicum gene #46 encoding metabolic pathway protein.  
 XX  
 KW Metabolic pathway protein; MP; lysine biosynthesis pathway;  
 KW methionine biosynthesis pathway; large-scale production of fine chemical;  
 KW *Corynebacterium diptheriae*; *diphtheria*; ds.  
 XX  
 OS *Corynebacterium glutamicum*.  
 XX  
 PN MO200166573-A2.  
 XX  
 PD 13-SEP-2001.  
 XX  
 PF 22-DEC-2000; 2000WO-IB02035.  
 XX







ID AAH65801 standard; DNA: 1422 BP.  
 XX AAH65801;  
 AC  
 XX  
 XX 26-SEP-2001 (first entry)  
 DT  
 XX C glutamicum coding sequence fragment SEQ ID NO: 836.  
 DE  
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
 KW organic acid synthesis; ds.  
 XX Corynebacterium glutamicum.  
 OS  
 XX EP1108790-A2.  
 XX  
 XX 20-JUN-2001.  
 PD  
 XX  
 XX 18-DEC-2000; 2000EP-0127688.  
 PF  
 XX 16-DEC-1999; 99JP-0377484.  
 PR 07-APR-2000; 2000JP-0159162.  
 PR 03-AUG-2000; 2000JP-0280988.  
 XX  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 PA  
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 PI  
 XX WPI: 2001-376931/40.  
 DR P-PSDB: AAG90582.  
 XX  
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analysing  
 PT expression profile or pattern of a gene and identifying homologous gene  
 PS  
 XX Claim 8; SEQ ID NO: 836; 246bp + Sequence Listing; English.  
 XX  
 XX The present invention provides a number of nucleotide and protein  
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of coryneform bacterium, measuring expression amount and  
 CC analysing the expression profile or expression pattern of a gene derived  
 CC from Coryneform bacterium, and identifying a homologue of a gene derived  
 CC from coryneform bacterium. Coryneform bacteria are useful for producing  
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a nucleic acid described  
 CC in the exemplification of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC European Patent Office.  
 CC  
 XX  
 XX Sequence 1422 BP; 309 A; 414 C; 401 G; 298 T; 0 other;

Query Match 40.7%; Score 594; DB 22; Length 1422;  
 Best Local Similarity 64.5%; Pred. NO. 3.5e-165;  
 Matches 920; Conservative 0; Mismatches 500; Indels 6; Gaps 2;

QY 32 tcgagacgaatgcgcgaatacactccatgtcttcgcgcgaaggaactaaccttg 91  
 DB 2 tggactcaaggtgcgcgaactcctcaacagagagcgagcagctacacagatcgtctg 61  
 QY 92 ctgaagaagaatgcgcgcgtctatgttctcgtgagcgtatcgcgtcttaagcat 151  
 DB 62 cagaatgatgatgacgcgcgtctcatgctggtcgcaagaattcgcaagcagcgacct 121  
 QY 152 tgaagggtgcgaatctcgttcctcccaatgacagctccagagcagcgctccatc 211  
 DB 122 tgaagggtgcgcgcgtctcgttcctcccaatgacagctccagagcagcgctccatc 181  
 QY 212 agacactcacagctcgtgctgatgacagatgagctccctcaacatctctctacac 271

DB 182 agaccctcactgcttggcgcgctgaggttcgttggccttcctcgaacatcttctccacc 241  
 QY 272 aagatacagcgctcgtctgtatctgttgcggcccaacagcagcagagagccagcg 331  
 DB 242 aggatgagcgtcgcagcgcgtctatcgttctgcgc--tcggcgacgtctgaagagcagctg 298  
 QY 332 gtatcccaactcgcctcgtgaaggcgcaaacactcccaatactctgggaacaacatacc 391  
 DB 299 ggttcacagatctcgtcgttggaaggtgacactgagagtaactggtgtgcataccacc 358  
 QY 392 ggcctcacaatgagcagatgttcaagggcccaacagcaggtgtgtgtagtggtagt 451  
 DB 359 agatctcagcctggggcgatgagct--gccaacatgaccccttcgaacagcggtgtgacg 415  
 QY 452 ctacactcctcactcccaagggcgttcgaatccgaacacgcgcgtgtcttcacagacca 511  
 DB 416 ccaacatgctgttattctgcgcgtcgcgaatacgaagcagcgtcgtctgtccacagcag 475  
 QY 512 cagaagctgacaacctcgaataccgcgtcgtcttctgaacatacgaagcttccaacc 571  
 DB 476 aggcacaagatctcgaatgatacgcacatctcttggcgtctgtgagttctgtctg 535  
 QY 572 aagacagaagaacacatgagcagagttcgtccgcgcgtgacgtgttcgaagagaca 631  
 DB 536 cagaagctgcaagtgaggcagaatcgtcgtgagcgttgaaggtgtcgcagcaggaaca 595  
 QY 632 caacaggtctcaccgcctcctacacagctcgtgaagaaggagggcaactcctcctccagca 691  
 DB 596 ccaacggtgtgacgcgcgtcttaccactccttggaagaggtgctgtcttccacagca 655  
 QY 692 tcaagctcaacagcagctgttacaagaagttccgaatacctcagcgttcgcgcacat 751  
 DB 656 tgaacgtcaacagcagcgtctgtacccaagttctgtaaaagaatgacgcgcacacat 715  
 QY 752 cccttatcgtatgatacaccgtgtctccgaatgcatatgacgtcgcgcagagacgtctcg 811  
 DB 716 cccgtatgcgcgcgcacataacacgcgcacatgcatatgagcgcgcagagacgtctcg 775  
 QY 812 tcatggttaacgagatgtcggcaaggctcgcgtcactccctcgttggcagagcgctc 871  
 DB 776 tctgcgttaccgagatgctgcggcaaggctcgtcgtcgtcgttccagcgcggcgcc 835  
 QY 872 gcttctcatcacagaactcgaacacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 931  
 DB 836 ggcgtcaaggtcaccagagcttcaacacacacacacacacacacacacacacacacacac 895  
 QY 932 aggtcgc 991  
 DB 896 ctgtgtcaccgcgttgatgagcgcacatcgagagcgcgcgcgcgcgcgcgcgcgcgcgcgc 955  
 QY 992 actgagatcatcctcgttgcacatgagtcgcccagatgagatgagatgagatgagatgagat 1051  
 DB 956 acaagagacatcattcctcctcgaagcagatgctcaagatgagatgagatgagatgagatgagat 1015  
 QY 1052 acatcgcacactcgtatgaacgaatltgatacagatgagcctcactgaataaccagacatca 1111  
 DB 1016 acatcgcgtatcgtatgaatgatgatgatgatgatgatgatgatgatgatgatgatgatgat 1075  
 QY 1112 agcacatcccaatcaagcagagatcagatgttgggaattcccaagttgcccagcgtatcc 1171  
 DB 1076 cccgcacacacatcaagcagcagatgagatgagatgagatgagatgagatgagatgagatgagat 1135  
 QY 1172 tccctctgtcgtgagggcgccgtcttcaactgtgctgctcagatgagatgagatgagatgagat 1231  
 DB 1136 tgcgt 1195  
 QY 1232 ttatgtcaatltcactcaacacagacacacacacacacacacacacacacacacacacacag 1291  
 DB 1196 tcatgtccaactccttcgcgcgtcagacacacacacacacacacacacacacacacacacacag 1255  
 QY 1292 gaatctcgtgagaagaagtttcaacacacacacacacacacacacacacacacacacacag 1351  
 DB 1256 gacagtaagagaagagttac 1315





[illegible]



PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137722.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139455.  
 PR 17-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139753.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140961.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142305.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142820.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 28-JUL-1999; 99US-0145919.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 05-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148345.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151303.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159299.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 29-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 34.8%; Score 508.8; DB 21; Length 1865;  
 Best Local Similarity 62.0%; Pred. No. 6.8e-140;





PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139464.  
PR 21-JUN-1999; 99US-0139750.  
PR 22-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142890.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145293.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0152933.  
PR 13-SEP-1999; 99US-0153070.  
PR 15-SEP-1999; 99US-0153758.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155436.  
PR 24-SEP-1999; 99US-0155436.  
PR 28-SEP-1999; 99US-0156559.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.



PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130047.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 04-MAY-1999; 99US-0132048.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 14-JUN-1999; 99US-0138847.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 21-JUN-1999; 99US-0139763.  
 PR 22-JUN-1999; 99US-0139817.  
 PR 23-JUN-1999; 99US-0138899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 28-JUN-1999; 99US-0140695.  
 PR 29-JUN-1999; 99US-0140823.  
 PR 30-JUN-1999; 99US-0140991.  
 PR 01-JUL-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142055.  
 PR 08-JUL-1999; 99US-0142390.  
 PR 09-JUL-1999; 99US-0142803.  
 PR 12-JUL-1999; 99US-0142920.  
 PR 13-JUL-1999; 99US-0142977.  
 PR 14-JUL-1999; 99US-0143542.  
 PR 15-JUL-1999; 99US-0143624.  
 PR 16-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 20-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144532.  
 PR 21-JUL-1999; 99US-0144684.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145199.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 28-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 04-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 10-AUG-1999; 99US-0147935.  
 PR 11-AUG-1999; 99US-0148171.  
 PR 12-AUG-1999; 99US-0148319.  
 PR 13-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 16-AUG-1999; 99US-0148684.  
 PR 17-AUG-1999; 99US-0149368.  
 PR 18-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151303.  
 PR 01-SEP-1999; 99US-0151438.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 28-SEP-1999; 99US-0155659.  
 PR 29-SEP-1999; 99US-0156458.  
 PR 04-OCT-1999; 99US-0156596.  
 PR 05-OCT-1999; 99US-0157117.  
 PR 06-OCT-1999; 99US-0157753.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.

PR 13-OCT-1999: 99US-0158369.  
 PR 13-OCT-1999: 99US-0159293.  
 PR 13-OCT-1999: 99US-0159294.  
 PR 13-OCT-1999: 99US-0159295.  
 PR 14-OCT-1999: 99US-0159329.  
 PR 14-OCT-1999: 99US-0159330.  
 PR 14-OCT-1999: 99US-0159331.  
 PR 14-OCT-1999: 99US-0159637.  
 PR 14-OCT-1999: 99US-0159638.  
 PR 18-OCT-1999: 99US-0159584.  
 PR 21-OCT-1999: 99US-0160741.  
 PR 21-OCT-1999: 99US-0160767.  
 PR 21-OCT-1999: 99US-0160768.  
 PR 21-OCT-1999: 99US-0160770.  
 PR 21-OCT-1999: 99US-0160814.  
 PR 21-OCT-1999: 99US-0160815.  
 PR 21-OCT-1999: 99US-0160980.  
 PR 22-OCT-1999: 99US-0160981.  
 PR 22-OCT-1999: 99US-0160989.  
 PR 25-OCT-1999: 99US-0161404.  
 PR 25-OCT-1999: 99US-0161405.  
 PR 25-OCT-1999: 99US-0161406.  
 PR 26-OCT-1999: 99US-0161359.  
 PR 26-OCT-1999: 99US-0161360.  
 PR 26-OCT-1999: 99US-0161361.  
 PR 26-OCT-1999: 99US-0161920.  
 PR 28-OCT-1999: 99US-0161992.  
 PR 28-OCT-1999: 99US-0161993.  
 PR 29-OCT-1999: 99US-0162142.

Query Match 34.1%; Score 497.8; DB 21; Length 1728;  
 Best Local Similarity 61.9%; Pred. No. 1.2e-136;  
 Matches 898; Conservative 0; Mismatches 503; Indels 49; Gaps 5;

QY 34 gagtacaagaattgcccgaacatcaactcattgtctcgcgcgttaagaagaacttaacctgtgt 93  
 DB 104 gaatacaagaattgcccgaacatcaactcattgtctcgcgcgttaagaagaacttaacctgtgt 163  
 QY 94 gagaagaagaattgcccgaacatcaactcattgtctcgcgcgttaagaagaacttaacctgtgt 153  
 DB 164 gaagtgtgagatgcccgaacatcaactcattgtctcgcgcgttaagaagaacttaacctgtgt 223  
 QY 154 aaggtgtgagatgcccgaacatcaactcattgtctcgcgcgttaagaagaacttaacctgtgt 213  
 DB 224 aaggtgtgagatgcccgaacatcaactcattgtctcgcgcgttaagaagaacttaacctgtgt 283  
 QY 214 aactcaagaactcgt 273  
 DB 284 aactcaagaactcgt 343  
 QY 274 gatacaagaactcgt 333  
 DB 344 gatacaagaactcgt 378  
 QY 334 atcccaagttctgcgt 393  
 DB 379 --cggt 436  
 QY 394 gcttcacatggt 453  
 DB 437 gcttcacatggt 496  
 QY 454 acaactcctcaatcccaagggt 501  
 DB 497 acctcttctgt 556  
 QY 502 ccagagcccaagaagaattgt 561  
 DB 557 cctgtatctctactgt 616  
 QY 562 gttctcaaccaagaagaaccatgt 621  
 DB 617 ggtctcaagaattgt 676

QY 622 gaagagacaacaagaaggt 681  
 DB 677 gagaagaatacaccatgt 736  
 QY 682 ttcccaagcatcaacatgt 741  
 DB 737 ttccctgtcatcaacatgt 796  
 QY 742 tgcgcgcacatccctatgt 801  
 DB 797 tgcgcgcacatccctatgt 856  
 QY 802 acaagctcgt 861  
 DB 857 gttgt 916  
 QY 862 caagagcgt 921  
 DB 917 gctgt 976  
 QY 922 gaaggtcaccaggt 981  
 DB 977 gaaggtcaccaggt 1036  
 QY 982 tgcacaagaagaatgt 1041  
 DB 1037 accacccgt 1096  
 QY 1042 attgt 1101  
 DB 1097 attgt 1156  
 QY 1102 ccaggtcaccaggt 1161  
 DB 1157 cctgt 1216  
 QY 1162 caagct--atcctcctcctgt 1218  
 DB 1217 aaggtcgt 1276  
 QY 1219 caaccatcctgt 1278  
 DB 1277 caaccatcctgt 1336  
 QY 1279 t-----acgaagaagaagaagaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1332  
 DB 1337 tgaaga 1396  
 QY 1333 gatgaaga 1392  
 DB 1397 gatgaaga 1456  
 QY 1393 cagaagaaga--gt 1451  
 DB 1457 aaggaacacacatcgt 1516  
 QY 1452 ccgt 1516  
 DB 1517 caaggt 1526

RESULT 13  
 AAT44513  
 ID AAT44513 standard; cDNA: 1767 BP.

XX AAT44513;

DT 22-FEB-1997 (first entry)

DE Atparagus S-adenosyl-L-homocysteine hydrolase cDNA clone.

XX S-adenosyl-L-homocysteine hydrolase; SHH; promoter; monocot; dicot;

transgenic plant; disease resistance; pathogen resistance; ds.  
 XX Asparagus officinalis.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 26..1483  
 FT primer\_bind /tag= a  
 FT primer\_bind complement (281..303)  
 FT primer\_bind /tag= b  
 FT primer\_bind /note= "PCR-1 primer"  
 FT primer\_bind 722..744  
 FT primer\_bind /tag= c  
 FT primer\_bind /note= "PCR-2 primer"  
 XX  
 XX W09632488-A1.  
 XX 17-OCR-1996.  
 XX 10-APR-1996; 96WO-GB00882.  
 XX 10-APR-1995; 95GB-0007381.  
 XX (ZENEC) ZENEC LTD.  
 XX Draper J, Greenland AJ, Skipsey M, Warner S;  
 DR WPI: 1996-477138/47.  
 DR P-PSDB; AAM01458.  
 XX  
 XX S-adenosyl-L-homocysteine hydrolase promoter - used for driving  
 PT expression of effector genes, such as pathogen resistance genes, in  
 XX transgenic plants  
 XX  
 XX Example 1; Fig 1; 57pp; English.  
 XX  
 XX A cDNA clone (AAT44513) codes for the S-adenosyl-L-homocysteine  
 CC hydrolase (SHH) (AAM01458) of asparagus. A cDNA clone, designated  
 CC DB6, was isolated from an asparagus cDNA library by homology  
 CC searching, and the full-length sequence was obtained by library  
 CC screening using DB6 as probe. The asparagus SHH sequence was used  
 CC to obtain SHH genes from other plant species. Analysis of the  
 CC Arabidopsis thaliana SHH gene (AAT44515) revealed a 1849 bp promoter  
 CC (AAT44514) which can be used to drive expression of effector (e.g.  
 CC pathogen resistance) genes in transgenic dicot and monocot plants.  
 CC  
 XX Sequence 1767 BP; 426 A; 462 C; 459 G; 420 T; 0 other;  
 SQ

Query Match 33.5%; Score 489; DB 17; Length 1767;  
 Best Local Similarity 61.1%; Pred. No. 4.8e-134;  
 Matches 886; Conservative 0; Mismatches 515; Indels 48; Gaps 4;

DB 337 -----ctccgtcttcgcgggaaggtgagaccctccaaagatctgtgtgtgacccagcg 393  
 QY 393 cgtctacatgagcagaatgtgacagggcccaagaggttgatgagtgatgagtgatgc 452  
 DB 394 tgccttgactggggcccggtgtgctgcgaactatctgcatgacgagcgagcaac 453  
 QY 453 taactctatctccaagggtc-----cgaaatcgaaacagcggtgtgtgt 500  
 DB 454 caactcttgatcatatgaggggtggaagcgaggaagaagtacgagaagcgagggaagt 513  
 QY 501 tccagagcccaagcagaatctgacaccccgatctgcgtcttctgtacatcaagca 560  
 DB 514 gcccgatccgggtctctacacgaatgtgagttccagatcgctgcacatcatcaagga 573  
 QY 561 ggtcttaacccaagacaagacacactgtgacacagttgtcgcgcatgaaaggtgttc 620  
 DB 574 tgggtcctaaagttgagccccaagatcaggaagatgaaagatgagttgtcgtgtgtc 633  
 QY 621 cgaagagacaacacaggtgtgacacgctctacacagctcgaagaaggaggaacacct 680  
 DB 634 ggaagagacacacacagcggtgtgacagagcttaccagatgaggtacaaatccctct 693  
 QY 681 ctcccaagccatcaacgacacagcgtgttacaagatccaaagttcgatcaatcaccg 740  
 DB 694 ttccctgcatcaatgtcaatgtatctgcgcacaaagacaaagttgacatctgtatg 753  
 QY 741 ctgcgcacactccctatcgaatgtatcaacggtgtctcgaatgtacatgcagcgcaa 800  
 DB 754 atgcgcgacactctctcccgatgtgtgagagggccacatgtatgattgtgtgcgaa 813  
 QY 801 gacagctctgcgtatgtgacgaggtatgacgacgaagggctgcgtcaatccctcgtg 860  
 DB 814 gtttcagttgtctcggtgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 873  
 QY 861 ccaagggctgcgtatcatatcacaagaactcccaactctgcgtctccagctgtgcac 920  
 DB 874 ggtgtgtcccggtgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 933  
 QY 921 ggaaggtcaacaggtgc 980  
 DB 934 ggaaggtctccaggtctccacactccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 993  
 QY 981 atgcacaggaactgcgatcatcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1040  
 DB 994 caccacgggttaacagagacatcatcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1053  
 QY 1041 tatgtcgtatcatcgc 1100  
 DB 1054 catgtctgcacacatgtgtcatcttgacacagagattgtgacatgtgtgtgtgtgt 1113  
 QY 1101 cccgcgcac 1160  
 DB 1114 cccgcgcac 1173  
 QY 1161 ccac-----gcatcctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1217  
 DB 1174 caaacactgtgtatgt 1233  
 QY 1218 tcaaccatcttctgt 1277  
 DB 1234 tcaaccacagcttctgt 1293  
 QY 1278 ctacgaaag-----agaaagaaatctcgaagaaggtgttacaacactccgaagatct 1331  
 DB 1294 gtggaatgagaagcaagcggaagtgatgagaagaaggtgttcgtccccaagatct 1353  
 QY 1332 cgaatgaagaggt 1391  
 DB 1354 tgaatgaagaggt 1413  
 QY 1392 acagaagcaggt 1451





Db 1092 catgtttgacacatgtgtcactttgacacacgaatcgcacatgtgtgtctcgagacct 1151  
QY 1101 ccagagcatcagacacatcccaatcaagccagaatatagcacatgtgggaattcccaatgg 1160  
Db 1152 cccgtgtgtcagaagatcacaatataagcctcaacacgaatgggtcttccctgacac 1211  
QY 1161 ccac----gctatcctccttctgtcgagagccgctcttaaccttggctgctgctacagg 1217  
Db 1212 caacagltggcatcaltgtcttggctgaagggtcgtcactgaacttggatgtgccaagg 1271  
QY 1218 tcaccatcttgcgtatgtacatgtcaltcacaacacgaacactcgctcagctcgacct 1277  
Db 1272 acaccctagtgttgtgagtgctgtcgctgtcacttaacaaagtcaticcccaactcgagt 1331  
QY 1278 ct-----acgaagaagagaagaatctcgaagaagaagtttacacacttccgaagcatct 1331  
Db 1332 gtggatgaaagaagcagtgagggaatgtgagaagaagtgtatgtcttggccaacacacct 1391  
QY 1332 cgaatgaagaagtcgctgcctcgaactcgatctctcgatgtccaccttaacaaagcttac 1391  
Db 1392 cgaacgaagaagltgtgcacttcaatcgcgaagaagctcggagccaagcttaaccaacttcc 1451  
QY 1392 acagaagcaggtcgaactacatcaacgttccagttgagggtccttacaagttcgtatgctta 1451  
Db 1452 gaaggtacaagctgaactatagcgttccagttgagggtccttacaagcctgtcacta 1511  
QY 1452 ccgttatata 1461  
Db 1512 caggtactga 1521

Search completed: August 17, 2002, 02:03:32  
Job time: 1131 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:48:22 ; Search time 2720.84 Seconds

(without alignments)  
11236.850 Million cell updates/sec

Title: US-09-759-990-1

Perfect score: 1461

Sequence: 1 atggcttgcgaatacaccctac.....ctgatgtcattacgtattaa 1461

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapept 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pal:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	1459.4	99.9	1461	6	AX195549	AX195549 Sequence
2	1440.2	98.6	1882	3	TYV40872	U40872 Trichomonas
3	655.2	44.8	37898	1	SCS34	AL35862 Streptomy
4	615.2	42.1	3800	1	SAR416377	AJ416377 Streptomy
5	596.4	40.8	1557	6	AX063937	AX063937 Sequence
6	596.4	40.8	1557	6	AX244105	AX244105 Sequence
7	596.4	40.8	1557	6	AX127145	AX127145 Sequence
8	594	40.7	1422	6	AX120920	AX120920 Sequence
9	535.4	36.6	1488	1	AP262755	AF262755 Mycobacte
10	535.4	36.6	1488	1	AX023852	AX023852 Sequence
11	535.4	36.6	18249	6	AE007145	AE007145 Mycobacte
12	535.4	36.6	13360	6	MYT20811	295121 Mycobacteri
13	529.6	36.2	1396	6	AX063941	AX063941 Sequence
14	529.6	36.2	1396	6	AX244109	AX244109 Sequence
15	512.6	35.1	1708	8	WHTSHH	LI1872 Trifolium ae
16	512.6	35.1	1837	8	MCU79766	U79766 Mesembryant
17	511.4	35.0	312050	1	MLEPRN3	AL583919 Mycobacte
18	506.6	34.7	1729	8	AY042866	AY042866 Arabidops
19	506.6	34.7	1802	8	AF059581	AF059581 Arabidops
20	506.6	34.7	1821	8	AF325037	AF325037 Arabidops
21	506.6	34.7	1821	8	AY049279	AY049279 Arabidops
22	506	34.6	13407	1	AE008877	AE008877 Agrobacte
23	506	34.6	13452	1	AE007946	AE007946 Agrobacte
24	501.8	34.3	1798	8	PSSADHY	X79905 Phalaenopsi
25	497	34.0	1790	8	ATSDLH	297059 Arabidopsis
26	490	33.5	1812	6	AR071283	AR071283 Sequence
27	490	33.5	1812	6	R03902	E03902 A gene invo
28	490	33.5	1812	6	TOBSALHH	D45204 Nicotiana t
29	489	33.5	1767	8	AS7643	AY57643 Sequence 1
30	488.2	33.4	215050	1	AL646057	AL646057 Ralstonia
31	483.6	33.1	2110	8	TOBCHP57A	LI6138 Tobacco MRN
32	481.6	33.0	1791	8	PUMSHB	M81885 Parsley S-R
33	475.6	32.6	10791	1	AE005699	AE005699 Caulobact
34	473	32.4	1685	8	AF428329	AF428329 Arabidops
35	472.4	32.3	1754	8	AF161705	AF161705 Lycopersi
36	471.4	32.3	1738	8	AY059888	AY059888 Arabidops
37	469.8	32.2	1843	8	AY050783	AY050783 Arabidops
38	466.6	31.9	1697	8	CRSAH1	226881 C. roseus SA
39	465.6	31.9	1878	8	TOBCHP57B	LI6139 Tobacco MRN
40	463.4	31.7	1824	8	AF185635	AF185635 Lupinus l
41	462.8	31.7	1824	8	ALFMSA25	L36119 Medicago sa
42	462.8	31.7	332635	1	AP003005	AP003005 Mesorhizo
43	459.8	31.5	11711	1	AE003941	AE003941 Xylella f
44	459	31.4	1660	8	AF462828	AF462828 Arabidops
45	455.8	31.2	10481	1	AE009636	AE009636 Brucella

#### ALIGNMENTS

RESULT 1	AX195549	1461 bp	DNA	linear	PAT 28-AUG-2001
LOCUS	AX195549	Sequence 1 from Patent WO0151651.			
DEFINITION	AX195549				
ACCESSION	AX195549				
VERSION	AX195549.1	GI:15386074			
KEYWORDS					
SOURCE					
ORGANISM					
	Trichomonas vaginalis.				
	Trichomonas vaginalis				
	Eukaryota; Parabasalidea; Trichomonadida; Trichomonadidae;				
	Trichomonas.				
REFERENCE	1 (bases 1 to 1461)				
AUTHORS	Hoffman, R.M., Xu, M., and Han, Q.				
TITLE	High expression and production of high-specific activity recombinant s-adenosyl homocysteine (sahh) and improved assays for s-adenosylmethionine (sam) Patent: WO 0151651-A 1 19-JUL-2001;				
JOURNAL	Anticancer, Inc. (US)				
FEATURES	Location/Qualifiers				
SOURCE	1..1461				
	/organism="Trichomonas vaginalis"				
	/db_xref="taxon:5722"				

BASE COUNT	372 a	427 c	335 g	327 t
------------	-------	-------	-------	-------

## ORIGIN

Query Match 99.9%; Score 1459.4; DB 6; Length 1461;  
 Best local Similarity 99.9%; Pred. No. 0;  
 Matches 1460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggttgcaataactactactggtgtgtcattgtgtatcagagattggccgaatcaaccc 60  
 DB 1 ATGGCTTGCAATACCTACTGCTGCTCTCATTCGATGAGATAGCAATGCGACATCAACCTC 60  
 QY 61 catgttcgagccgtaaggaactaaccttgcctgagaaagaaatgccaggtctatgatt 120  
 DB 61 CATGTTCTGGCCGTAAGGAATTAACCTTGCTGAGAAAGAAATGCGACAGTCTTATAGTT 120  
 QY 121 ctcttgagagcgtatccgcttcttaagccaatgtaaggtgtcaagaatctctgttccctc 180  
 DB 121 CTTCGTGAGCGTATTCCTGCTTCTAAGCCATTGAAGGGGTGTGAGAAATCTCTGGTCCCTC 180  
 QY 181 caatagaactcagagacagccgtctctatcagacactcacagccttgggtgtatgtc 240  
 DB 181 CACATGACAGTCCAGACAGCCGCTCTCATTCGAGACATCACACCTTTGGTGTGATGTC 240  
 QY 241 agatggccttctctgaacatcttctctacacaagaatcacagccgctgtcgtatcgttgc 300  
 DB 241 AGATGGGCTTCTGCAACATCTTCTCTACACAAGATACAGCCGCTGCTGTATCGTTGTC 300  
 QY 301 ggcaccaagagacacacagagagagccagcgtatccagcttgccttggagagggagaa 360  
 DB 301 GGCCCAACAGGACACACAGAGAGCCAGCGGTATCCAGTCTTCGCTGGAAGGGGAGAA 360  
 QY 361 aactcaccagaatacttggagagacacatcacccgctctcacatgagcagagtggtlcaag 420  
 DB 361 AACTCTCCAGAAATACCTGGAGAGAACATACCCGCTCTCACATGGCCAGATGGTCAAGGC 420  
 QY 421 ccacagagagttgtcagatgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480  
 DB 421 CCACAGCAGGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 QY 481 ttggaagaagcgggt 540  
 DB 481 TTGGAAGAAGCGGCTGT 540  
 QY 541 gtcttctgactactcaagcaggttctcaacccaagaagaacacactgtgcaacagttgct 600  
 DB 541 GTTCTGTCTACTACTCAAGCAGGTTCTCAACCAAGAACACCACTGGCCACAGTGTCT 600  
 QY 601 gccggatgaaacggtgttctcgaaagagaaacagaggtgttccaccccttaccagctc 660  
 DB 601 GCCGGCATGAACGGTGTTCGAAAGAACAAACAGGTGTCCAACCCCTCTTACCAGCTC 660  
 QY 661 gagaagaggggcaaacctctcttccagcactcaacagctcaacagcgtgttacaagaatcc 720  
 DB 661 GAGAAAGAGGGCAAACTCTCTTCCAGCATCAACGTCAACGACGTGTATCAAAAGTCC 720  
 QY 721 aagttcgataactaactaagcgtgcgcacatcccttatcgatggtatcaacagcgttcc 780  
 DB 721 AAGTTGATTAACATCTACCGCTGCCGCCACCTCTTATCATGATGATGATCAACCGTGTCC 780  
 QY 781 gatgcatgatcggcgcaagaagcgttctgtcatggtttagcagcgtatgttgcgaagggc 840  
 DB 781 GATGTCATGATCGCGGCAAGACAGCTCTGTCTATGAGGTTACGGCGATGTGCGCAAGGCG 840  
 QY 841 tgcgctcaatccctcgttgcgaagcagcgtcgttatacatcacaagaactcgaccacatc 900  
 DB 841 TGCGCTCAATCCCTCGTGGCCAAAGCGCTGCGTATCATCAAGAAATGACCAATTC 900  
 QY 901 tgcgcttccagcgtcgtcagatgaaagcgttaccaggttccgcatcggaggaagtgtcag 960  
 DB 901 TGCGCTTCCAGCGTCCATGAAAGGTACAGGTTCCGCGCATGAGGAAGTGTGTCAAG 960  
 QY 961 gatgcatatctgttcatatcagagaaactgcatatcatctgttgaatagtg 1020

DB 961 GATGTCATATCTTGTTACATGCAACAGGAACCTGCGATATCATCTCTGTGACATGATG 1020  
 QY 1021 gccagatgaaagagataagcgtatgttcgttaacatgcgcaccttcgataacgaattgat 1080  
 DB 1021 GCCAGATGAAGAGATGAAGCTATTTGTCGTAATGCGCCACTTCCATTAACGAATTGAT 1080  
 QY 1081 acagatgacctatgaatataccacagacatcaagcacatcccaatcaagcagaatagcag 1140  
 DB 1081 ACAGATGCGCTTCATGAATAATACCAGGACATCAAGCACATCCCAATCAAGCCAGATACGAC 1140  
 QY 1141 atgtggagatcccgatgagcagcgtatccctctcttctgttagggcgcccttcaac 1200  
 DB 1141 ATGTGGATTTCCAGATGCGCACGCTATCTCTCTTCTGTAGGGCCGCTCTTAAAC 1200  
 QY 1201 ctgtgctgcgtacaggttaccacatcttctgtatgtlcaatgtcatcacaacacagaca 1260  
 DB 1201 CTGTGCTGCGCTACAGGTACACCATCTTTCGTTATGATGATGATGATGATGATGATGAT 1260  
 QY 1261 ctgcgtcagcttgcacctacagaagaagagaaatctcgagaagaaggttatacactt 1320  
 DB 1261 CTGCTCAGCTGACCTCTACGAAAGAGAGAAATCTCGAAGAAAGGTTTACACACTT 1320  
 QY 1321 ccgaagcgtctgataagaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1380  
 DB 1321 CCGAAGCATCTGATGAAGAAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 QY 1381 acaagcgtacacagaagcagcgtacatacaacagcttccagttgaggttcctacaag 1440  
 DB 1381 ACAAAGCTTACACAGAAAGAGCGGTGACTATCATCAAGCTTCCAGTTGAGGGCTTACAA 1440  
 QY 1441 tctgagcttaccgtatataa 1461  
 DB 1441 TCTGATGCTTACCGTTATTAA 1461

RESULT 2  
 TVU40872 1882 bp DNA linear INV 31-OCT-1996  
 LOCUS Trichomonas vaginalis S-adenosyl-L-homocysteine hydrolase gene,  
 DEFINITION complete cds.  
 ACCESSION U40872  
 VERSION U40872.1 GI:1118008  
 KEYWORDS Trichomonas vaginalis.  
 SOURCE Trichomonas vaginalis.  
 ORGANISM Eukaryota; Parabasalidea; Trichomonadida; Trichomonadidae;  
 Trichomonas.  
 REFERENCE 1 (bases 1 to 1882)  
 AUTHORS Bagnara,A.S., Tucker,V.E., Minotto,L., Howes,E.R., Ko,G.A.,  
 Edwards,M.R. and Daves,I.W.  
 TITLE Molecular characterisation of adenosylhomocysteine from  
 Trichomonas vaginalis  
 JOURNAL Mol. Biochem. Parasitol. 81 (1), 1-11 (1996)  
 MEDLINE 97047381  
 REFERENCE 2 (bases 1 to 1882)  
 AUTHORS Bagnara,A.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-1995) Aldo S. Bagnara, Biochemistry and Molecular  
 Genetics, The University of New South Wales, Sydney, New South  
 Wales, 2052, Australia  
 FEATURES  
 source Location/Qualifiers  
 1..1882  
 /organism="Trichomonas vaginalis"  
 /strain="WA38"  
 /specific\_host="Homo sapiens"  
 /db\_xref="taxon:5722"  
 1..290  
 291..1751  
 /ec\_number="3.3.1.1"  
 /codon\_start=1  
 /product="S-adenosyl-L-homocysteine hydrolase"  
 /protein\_id="AAC47319.1"  
 /db\_xref="GI:1118009"

/translation="NACKSPAGAPPEYRIADILUHLVGRKELTAEKEMPGLVLRER  
 TSASPLKGVRLSSLIHMTVQTALETTLTALGADVMASCNFTSTDTAAALIVGP  
 TGTPEKPAIPEVANKGELTPEWENTYVGLTPDSCGPOOVYDGDATLLISKFE  
 FETAGAVEPEPTADNLEKCVLATLEKOVNODNHHMTVAAGNVSEPTTGVHRLY  
 OLEKGLLEPAIINVDATKSPDNVIGCRSLDINGINASNVIGKRALVNGYD  
 VKGCAOSLROGARYITTEPDIOLQVMESTQVRIEYKADVFETCTGNCDI  
 ISVDMQMKDRAITVGNITHNEIDTDGLMKPGIKHIDIKEDWMEPPDGHAILL  
 LAEGRLNIGCATGHPSEFVMSFTNQLAIDLKRGMLEMKVYTLPRHLEEVYR  
 LHLGSLDHLTPTLTKOKROADYINVPYEGPYKSDAYRY"  
 1752..1882

3'UTR  
 polyA-signal  
 BASE COUNT 506 a 499 c 395 g 482 t  
 ORIGIN

Query Match 98.6%; Score 1440.2; DB 3; Length 1882;  
 Best Local Similarity 99.1%; Pred. No. 0;  
 Matches 1448; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 atgcttgcgaatacactactgtgtgtcattcagatagacgaattgcgaatacactc 60  
 Db 291 ATGGCTTCAATACACCTGCTGCTCATTGAGTACAGAAATGCGCATCAACCTC 350  
 QY 61 catgttcgcgcgaagaacttacccttgcctgagaagaattgcgaattcattagt 120  
 Db 351 CATGTCTCGGCGCTTACAGAACTTACCTTGTGAGAAAGAAATGCCAGGTCTTATG 410  
 QY 121 ctctgtgagcgttattcgcctcctaagcattgaagggtgtcagaatctctgttc 180  
 Db 411 CTCTGTGAGCCTTATTCCGCTTCAAGCATGAAAGGTGTGAGAAATCTCGTTCCTC 470  
 QY 181 cacatgacatccagaagcgcctcctcctcagacactcaaaccttctgtgtcgtc 240  
 Db 471 CACATGACAGTCCACACAGCGGTCTTATTGAGACACTCACACTCTTGGTGTGAGT 530  
 QY 241 agatggccttcctgcgaatacttctcacaagaatacagcgcctgtctgtatc 300  
 Db 531 AGATGGGCTTCTGCAACATCTTCTTACAAAGATACAGCCCTGCTGCTATCTGTC 590  
 QY 301 ggcgaagaagcacaacagaagaagcagcgttaccagcttctgcgttgaaagg 360  
 Db 591 GGCCCAACAGCAGCACACAGAGAACAGCGGTATCCAGTCTTCCGCTGGAAGGGG 650  
 QY 361 aactccagaatacttggaagaacacatacgcgcctcctacatggcgaagtgtc 420  
 Db 651 AACTCCAGAAATCTGGAGACACATACCGCTCTCACATGGCCAGATGGTCAAGGC 710  
 QY 421 ccacagaagttgtcagatggtgtgtatgtacactcctcctcacaaggcctcga 480  
 Db 711 CCACAGCAGGTTGTGATGATGATGTGTATGCTTACACTCTCATCTCCAAAGGCTT 770  
 QY 481 ttggaacagcgcgtgtgtcttcagaagaacagaagctgaacactcgaatac 540  
 Db 771 TTGGAACAGCGGTGTGCTCCAGAGCAACAGAGGTGAGCAACTCTGCAATACCG 830  
 QY 541 gtcttctcactcaagcaggtcttcaacagaagaacacagcagcagcaagttgt 600  
 Db 831 GTTCTTCTACACTCAACAGCAGGTCTTCAACAGCAACAAACACAGTGGCAAGT 890  
 QY 601 gccggaatgaagtglttcgaagaagaacaacagtgltccacgcctctaccag 660  
 Db 891 GCCGGCATGAAAGGTGTTCCGAAGAGACAACACAGGTGTCCACCGCTCTACCA 950  
 QY 661 gagaagaagggaacacctcttccagccatcaacgctcaagaagcgtttacaag 720  
 Db 951 GAAAGAGAGGCAAACTCTTCCCAACCAACAGCAACAGCAAGCTGTTCACAAAGT 1010  
 QY 721 aagttcgaatacactcagcgttcgcacacccctcctcagatgattcaacgct 780  
 Db 1011 AAGTTGATTAACATCTACGCGTGTGCGCACCTTATCGATGATCAACCGCTTCC 1070  
 QY 781 gatgtcatgattcggcgaagaagcctctcgtcatgggttaccgagatgtcgcaagg 840

Db 1071 GATGTCAATGATCGGCGCAACAGACCTCTCGTCAATGGGTTACGGCGATCGGGAAGGC 1130  
 QY 841 tgggtcaatccctccgttggcgaaggcgcgtgtatcattacagaactcgaacac 900  
 Db 1131 TGGGCTCAATCCCTCCGCGGCAAGGCGCTCGCTTATCATCAAGAACTCGACCTTATC 1190  
 QY 901 tggcctcgaagctgcattgaaaggaagcagcgcgcgaatcgaagaagtcgcaag 960  
 Db 1191 TGGGCTCTCAAGCTGTGATGAAAGGCTTACAGAGGTCGCGGATGAGGAAGTCTCAAG 1250  
 QY 961 gatgtcatatcttcgttacaatgcagaagaactgcgatalatcctctgttgcatt 1020  
 Db 1251 GATGTGATATCTTCTTACATGCAACAGAACTCGATATCATCTTGTACATGATG 1310  
 QY 1021 gcccaatgagaataagcgtattgtgtgtatcatcgcgcacctcgtataagaattgat 1080  
 Db 1311 GCCCAGATGAAGGATTAAGGCTATTGTGCGTAAATCGGCGCATCTTGATTAACAAAT 1370  
 QY 1081 acagatggcctcattgaataaccaggaacacacacacacacacacacacacacac 1140  
 Db 1371 ACAGATGGCCTCATGAATATACCGGCAACGATCAAGCATCCCAATCAAGCAGAAAT 1430  
 QY 1141 atgtggaattcccaatgagcagcgtatcctcctctgtgtaggcgcgcctctaac 1200  
 Db 1431 ATGTGGAAATTCACAGATGGCCACGCTATCTCTTGTGTGAGGGCGCCTTCTTAC 1490  
 QY 1201 ctgtgctgcctcagatgacacacacacacacacacacacacacacacacacacac 1260  
 Db 1491 CTGTGTTGGCTTACAGTACCCATCTTCTTGTATGATCAATGATCAACAAACAGAA 1550  
 QY 1261 ctgcctcagctcgcacctcctcgaagaagaagaatcctcgaagaagaagttacac 1320  
 Db 1551 CTGCTCAGCTCAGCTCTACGAAAGAGAGAAATCTCGAATGAAGGTATTACACACT 1610  
 QY 1321 ccgaacatctcgtatgaagaagtcgtcgcctcctcagctcgtatctcgtatc 1380  
 Db 1611 CCGAAGCATCTCGATGAAGAGTGTTCGCTCCACCTCGGATCTCTGATGTCCACTT 1670  
 QY 1381 acaaagcttacacagaagcagcgtgactacatacagcttcaagttgaagttcctaac 1440  
 Db 1671 ACAAAGCTTACACAGAAAGAGGCTGACTATCAACGTTCCAGTGAAGGTCCTTAACAG 1730  
 QY 1441 tctgattctaccgttattaa 1461  
 Db 1731 TCTGATGCTTACCGTTATTAA 1751

RESULT 3  
 SCE34/c  
 LOCUS 37898 bp DNA linear BCT 01-JUN-2000  
 DEFINITION Streptomyces coelicolor cosmid E34.  
 ACCESSION AL353862  
 VERSION 1  
 KEYWORDS adenosylhomocysteine; integral membrane protein; L-lactate permease; mannose-6-phosphatase isomerase; nucleotide phosphatase; oxidoreductase; phosphomannomutase; rmd; secreted protein; transcriptional regulator; transport protein; whb.  
 Streptomyces coelicolor A3(2)  
 SOURCE  
 ORGANISM  
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 1 (bases 1 to 37898)  
 Redenbach, M., Kleser, H.M., Denapate, D., Eichner, A., Cullum, J., Kinsahl, H. and Hopwood, D.A.  
 A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome  
 Mol. Microbiol. 21 (1), 77-96 (1996)  
 JOURNAL MEDLINE 97000351  
 2 (bases 1 to 37898)  
 REFERENCES Saunders, D.C. and Harris, D.  
 JOURNAL Unpublished  
 3 (bases 1 to 37898)  
 Cerdano, A.M., Parkhill, J., Barrell, B.G. and Randal, M.A.

TITLE  
JOURNAL

## COMMENT

Direct Submission  
Submitted (25-APR-2000) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Cosmids Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK

Notes:  
Streptomyces coelicolor sequencing at the Sanger Centre is funded by the BBSRC and Beowulf Genomics  
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.  
(URL: <http://www.sanger.ac.uk/projects/s.coelicolor/>)  
CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand)  
The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.  
The length in codons is given for each CDS.  
Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon.  
Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.  
IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.  
Cosmid E34 overlaps cosmid E33 on the AseI-E genomic restriction fragment.

## FEATURES

## source

## misc\_feature

## rRNA

## rRNA

## rRNA

## gene

## CDS

1. 37898  
/organism="Streptomyces coelicolor A3(2)"  
/strain="A3(2)"  
/db\_xref="taxon:100226"  
/clone="cosmid E34"  
1. 161  
/note="nominal overlap with Streptomyces coelicolor cosmid StE33"  
complement(1..105)  
/note="probable 23S rRNA (fragment), coordinates taken from similarity to SW:SARRND (EMBL:M27245) Streptomyces ambifaciens rrrd gene cluster encoding 16S, 23S and 5S ribosomal RNAs"  
complement(1..2116)  
/note="coordinates taken from similarity to SW:SARRND (EMBL:M27245) Streptomyces ambifaciens rrrd gene cluster encoding 16S, 23S and 5S ribosomal RNAs"  
complement(2470..3477)  
/gene="SCE34.01c"  
complement(2470..3477)  
/gene="SCE34.01c"  
/note="SCE34.01c, possible integral membrane protein, len: 335 aa; similar to TR:069662 (EMBL:AL022121) Mycobacterium tuberculosis hypothetical 35.5 kD protein MTW025.042c, 330 aa; fasta scores: opt: 1119 z-score: 1285.3 E(): 0; 51.8% identity in 334 aa overlap. Contains possible hydrophobic membrane spanning regions"

RBS  
RBS  
CDSgene  
CDSgene  
CDS

/codon\_start=1  
/transl\_table=11  
/product="putative integral membrane protein"  
/protein\_id="CAB88904.1"  
/db\_xref="GI:7649484"  
/translation="MDLDVFSAHRAEMDRDLARRRRITGPEDVLAVALQRTAT  
HLSLIQSSAPDQTLGRSLQVLAARASVTGTRASRDVTRLTQGFPAVYRARHM  
WPTALISTAVAAALGLGVLGVLPIWIFQNLNAGVGLMSAGRLDTPLGLVLP  
AQQVNTNMAALCLILGVLGVLPIWIFQNLNAGVGLMSAGRLDTPLGLVLP  
HGLELTFVFAAGTGLRGLGLIDGPPTRRLALEGFAAIGMAVGLALVFGCA  
TEGVTPLGTPWARTITGVLALFLAVYVYVIGSAVRGEGVDGAARASATVPA  
A"  
complement(3489..3494)  
3589..3595  
3604..4560  
/gene="SCE34.02"  
3604..4560  
/gene="SCE34.02"  
/note="SCE34.02, possible integral membrane protein, len: 318 aa; similar to TR:069663 (EMBL:AL022121) Mycobacterium tuberculosis hypothetical 33.2 kD protein MTW025.043, 310 aa; fasta scores: opt: 734 z-score: 602.4 E(): 4.3e-26; 42.9% identity in 319 aa overlap. Contains possible hydrophobic membrane spanning regions"  
/codon\_start=1  
/transl\_table=11  
/product="putative integral membrane protein"  
/protein\_id="CAB88905.1"  
/db\_xref="GI:7649485"  
/translation="MSEVTGSEVAVLELRPARLPSRALAVLDLAVAVVAVTIAL  
MYATISLDVAQTAISATFVLYVGVPTAVELTSHRSIGKMACGRVYDDGGPIR  
FRHSIVRGITGYITLMTGLVACVLSVARSRRSDVTRAGTIVYERPFSSAGFM  
PPEPPLAGREFGLDLSAVDDLMILAVROVLARMQGLDPVGMMAARLADVAARTG  
APVPEVPPPAYLAVALQEQARARARAFEGASAEGTAAAVPAVPAVPAVPAV  
PVPVPAVPAVPAVPPSPGGLVPPVDPVDRRRGTGFVPA"  
complement(4581..5195)  
/gene="SCE34.03c"  
complement(4581..5195)  
/gene="SCE34.03c"  
/note="SCE34.03c, unknown, len: 204 aa"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein SCE34.03c"  
/protein\_id="CAB88906.1"  
/db\_xref="GI:7649486"  
/translation="MPGRGSLDHPDHPDLAEHFOCAPSGSGRRYSRLTTPGDA  
QGSVDLAIDELGRPIRLLEMAAMQVGAALDGTWTRPTGTSLHAYEGNARAFATG  
TSPAFVAVARLRLTPSEATRVRLVALDPVLAPEPTVQSMALLNREAHATNDPL  
TVDEYQVTLADTGEQYAVHAGDVVLAAPGLELDELTPEPSVFP"  
complement(5306..6763)  
/gene="sahH"  
/note="SCE34.04c"  
complement(5306..6763)  
/gene="sahH"  
/note="SCE34.04c, sahH, adenosylhomocysteinase, len: 485 aa; highly similar to SW:SAHH\_TRIYA (EMBL:U40872) "Rhizomys vaginalis adenosylhomocysteinase (EC 3.3.1.1)"; 65.2% identity in 477 aa overlap. Contains Pfam match to entry PF00670 AdoHcyase, S-adenosyl-L-homocysteine hydrolase and matches to Prosite entries PS00738 S-adenosyl-L-homocysteine hydrolase signature 1 and PS00739 S-adenosyl-L-homocysteine hydrolase signature 2"  
/codon\_start=1  
/transl\_table=11  
/product="adenosylhomocysteinase"  
/protein\_id="CAB88907.1"  
/db\_xref="GI:7649487"  
/translation="MTTVDRKDPKVDLSLAARGRKITLAHEHMPGLMAIRKRYAE  
AQLAARVYGSAMVYQTVALLFTVALGAQVMAVMAVMAVMAVMAVMAVMAV  
PDPPQGVPAFAMGRTLEETWCKTEQALTPNPTGCGNMLDGDGAATLVHAKVEY  
EDKGVPSVTAASDEHRLVLELTVGVSPQWQTLASEIRGTEETTVGNHLYE  
MHRDGLTLPALVNDVAVTKSKFDNKKYCRSLIDGINRAVDVLLGKTAAGVGYDV

KGCAESI.RGCGARVITTEIDPICALQAMDFQVTTIDEVYKADIPVTTGNKDI  
MAKDAKKHQAIVGNIGNEDNEIDMAGIADQPGIADKPEKPOVHTTYPEDKVIYL  
SEGRLLINGNATGHSFVMSNFADDTLQIETLPTDEPDTVDVLPKHLDERVARL  
HDSLGKLVITLTPSOADITGVKESGVPADHYR"  
Complement(5480..6733)  
/gene="sahh"  
/note="Pfam match to entry PF00670 AdoHcyase,  
S-adenosyl-L-homocysteine hydrolase, score 1004.10,  
E-value 3.3e-298"  
Complement(5921..5971)  
/gene="sahh"  
/note="PS00739 S-adenosyl-L-homocysteine hydrolase  
signature 2"  
complement(6479..6523)  
misc\_feature  
/gene="sahh"  
/note="PS00738 S-adenosyl-L-homocysteine hydrolase  
signature 1"  
complement(6773..6778)  
RBS  
complement(7037..8017)  
/gene="SCE34.05c"  
/complement(7037..8017)  
CDS  
/gene="SCE34.05c"  
/note="SCE34.05c, possible transport protein, len: 326 aa;  
similar to TR:AF26917 (EMBL:AF210843) Sorangium  
cellulosum putative transport protein, 305 aa; fasta  
scores: opt: 629 z-score: 707.7 E(): 5.9e-32; 37.88  
identity in 299 aa overlap. Contains Pfam match to entry  
PF01545 Cation\_efflux, Cation efflux family and possible  
hydrophobic membrane spanning regions"  
/codon\_start=1

Query Match 44.8%; Score 655.2; DB 1; Length 37898;  
Best Local Similarity 66.6%; Pred. No. 2.9e-140;  
Matches 955; Conservative 0; Mismatches 473; Indels 6; Gaps 1;

QY 34 gagtacaagaattgcgaacatcaactcatgtctcgcgcgtaagaactcctgct 93  
Db 6739 GACTTCAAGTTCGCCGACCTTCCTCGCGCGGTTGGCGGCAAGATCACCTCGCC 6680  
QY 94 gagaagaagaatgcaggtccttctgttctgtgagcgttatctcgtcttaagcaatg 153  
Db 6679 GAGCAGCAGATGCGCGGCTCATGGGATCCGCAAGAGTACCGCGCCAGCCGCTC 6620  
QY 154 aagggtgcagaatcttggtccctccatgacatgacgttcgaagcgcgtctatgag 213  
Db 6619 GCCGGCGCCCGGTACCGGCTCCCTGCAATGACCGTGAAGACCGCGCTCATGAG 6560  
QY 214 aaactcaagctctgtgtgtgtatgacatggtctcctgcaacatctctctacaaa 273  
Db 6559 ACCCTGTCGCGCTCGCGGCGGAGGTCGCTGGGCTCTCGCAACATCTTCTCACCCAG 6500  
QY 274 gatacagcgcgtgtgtctatggttcggaacacagcgaacaggaagcagcgt 333  
Db 6499 GACCAACGCGGCGCGGCTATCCGCGGCAAGCGCCGCAACCCGCAAGGCG 6440  
QY 334 atccagctcttcgcctggaaggaagaacatccagaatctgggaacacataacgc 393  
Db 6439 GTCCCGCTTCGCTCGCTGGAAGGCGGACCTCTGAAGAGTACTGGTGGTCTACGAGACAG 6380  
QY 394 gctctcaatgycagatgltca-----ggccacagcaggttgttgcatagtatgct 447  
Db 6379 GCGGTGACCTGCGGCAACACCCGCGGCGGCGCAACATGATCTCGAGCGAGGGGT 6320  
QY 448 gatgtacatctctcatctcaaggtctggaattcgaacacagcgcgtgtcttcagag 507  
Db 6319 GACGCCACCTTCCTCGTCCACAAAGGCGCTGAGTACGAAGAAGCAGCAAGTCCCTCG 6260  
QY 508 ccaacagaagctgaacactgaacatccgtctgtctgtctacatcaagcaggtcttc 567  
Db 6259 GTGCAACCGCGGAGTCCGACGACGACCGGTCATCTCGAAGTCTTACACCCGACGCTC 6200  
QY 568 aaccaagaagaacactggtcacacagttgtcgcgcgcatgaacggtgttcgaagag 627

Db 6199 GCGGAGAGCCCGCAGAACTGCACCCAGCTGGCGTCCAGATCCGCGGTCTCACCGAGAG 6140  
QY 628 acaacaacaggtgtccaccgcctctaccagctcgaagaaggaagcaactcctctccca 687  
Db 6139 ACCACGACGCGCGCTCACCCGCTTACAGATGACACCGGAGGACACCTTGCTTCCCG 6080  
QY 688 gcatcaagctcaacagcgcgtgttacaagaattccaaagttcagatacaatctagctgcgc 747  
Db 6079 GCGATCAACGTCAACGACGCGGTACCAAGTCAAGTTTCAGAACAAATACGCTGCCG 6020  
QY 748 cactcccttactgattatlaaacgcgtctcagatctgcatcattatggcggaagaagct 807  
Db 6019 CACTCCCTTGATTCAGCGGCATACCGGATACCGGCGAGTCTGATGCGGCAAGACCGCG 5960  
QY 808 ctgcgtcattcagcgcgtatgctgcaagaggtctgcgtcctcctccgttcggaagc 867  
Db 5959 GTCGTCTGCGGCTACGCGGAGAGGTCGCGCAAGGCGTCCGGGATCTCCCTGCGGCGGCG 5900  
QY 868 gtcgcggttatcattcaacaaactcgaacccaatctgcgtctcagcgttcagatgaaagc 927  
Db 5899 GCGCGGTGATCATCACCGAGATGACCCGATCTCCGCTGCAAGCGGCGGATGAGCGGC 5840  
QY 928 taccagctgcgcgcacgaagaagtgctaaagatgctcgtatcttgattatcagacaa 987  
Db 5839 TTCAGGTTCACACGCTGAGAGGTCGTGCAAGGCGGACATCTTCTTACACGACG 5780  
QY 988 ggaacatgcatactatctctgtgacatgagtgccagatgaagataagctatgtc 1047  
Db 5779 GGCACAAAGACATCATCATGCGGCAAGGACATGCGCAAGTATACACGACGCGCATGTC 5720  
QY 1048 ggttaacatgcgcacactcgaatacgaatgtatacagaatggtcctatgaataacagc 1107  
Db 5719 GGCACATCTCGGCGCATCTTCGACMACGAGATGACATGCGCGGCTGCGCACGACCCGCG 5660  
QY 1108 atcaagacatctcccaatccagcgaagaatgacatgttggaattccagatgccaagct 1167  
Db 5659 ATGCTCAAGACGACGATGCAAGGCCAGGTCTCACACTTGGACCTCCGACGCGGAGTGG 5600  
QY 1168 atccctctctgtcgaagcgcgctcttcaactctgctgcgtcgtacagatcaacatct 1227  
Db 5599 CTCATGCTCTTCGAGGCGCGCTCTGTGAACCTGGGCAAGCGGCGGACACCGCTTG 5540  
QY 1228 ttgcgtatgtcaatgtcatccaacacagacactcgtcgtcgtacgtcgaactcagaaga 1287  
Db 5539 TTCGTATGTCCAACTCTTCGCGGACGACACCTGGCGGATGAGGTTCATCAAG 5480  
QY 1288 agaggaatctcgaagaagaggtttacacacttccgaacatctcgaatgaagaatgct 1347  
Db 5479 CCGAGAGAGTACCCGACGACGCTTACGCTGCTGCCAAGCACCTTGAGAGAGTGGCG 5420  
QY 1348 cgcctcaactcgaatctcgtatgtccaccttcaaaagcttacaagaagaagcgtgac 1407  
Db 5419 CGGCTCACCTCGATCCCTTCGCGGTCAAGTGCACAGTCCGCGGAGGCGGCGAG 5360  
QY 1408 tacatcaagttcagttgaggttccttaagaagctgattacgttatataa 1461  
Db 5359 TACATCGGCTCAAGGTGAGGCGCTTACAAAGCGGACCACTACCGCTACTGA 5306

RESULT 4  
SAR416377 3800 bp DNA linear BCT 15-OCT-2001  
LOCUS SAR416377  
DEFINITION Streptomyces argillaceus mtnz gene, mtna gene and mtnh gene.  
ACCESSION AJ416377  
VERSION AJ416377.1 GI:16197721  
KEYWORDS adhcyase; AdoMet synthetase; MTHF reductase; mtna gene; mtnh gene;  
mtnz gene; thioesterase.  
SOURCE Streptomyces argillaceus.  
ORGANISM Streptomyces argillaceus  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
REFERENCE 1 (bases 1 to 3800)  
Fernandez Lozano,M., Lombó,F., Brana,A.F., Salas,J.A. and Mendez,C.

TITLE Generation of a Streptomyces argillaceus mutant by deleting two methyl cycle-related genes located within the mithramycin gene cluster leads to higher production of mithramycin

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3800)

AUTHORS Mendez, C.

TITLE Submitted

JOURNAL Submitted (05-OCT-2001) Mendez C., Biología Funcional e IUOPA, Universidad de Oviedo, C/ Julian Claveria s/n, 33006 Oviedo, SPAIN

FEATURES

source Location/Qualifiers

1.3800

gene /organism="Streptomyces argillaceus"

CDS /db\_xref="taxon:41951"

7..789

/gene="mtm2"

/function="putative role in polyketide chain biosynthesis and release"

/codon\_start=1

/transl\_table=11

/product="putative thioesterase"

/protein\_id="CAC94888.1"

/db\_xref="GI:16197722"

/translation="MAAOPITLXCFANAGVSAFHGMARVDPVVPVLLPGHD RRRREPVTTRAAALDDVRLFGAPTAPEVLVIGSLGALVATVGRALHEAGRGPA LLAAGCPPEPLPEPPAYALGPDALERTLLMAYVPSGTPEAGIIMDRVLPVLRDQLTARALACTSSPADLITPLILVADDDMAPAVAEWGSMSSEGTPLHIT VSGGHFVRGPELRLILGACALPARPRTPAAGVSTGN"

918..2300

gene /gene="mtm4"

CDS /db\_xref="taxon:41951"

918..2300

/gene="mtm4"

/function="putative role in an activated methyl cycle"

/codon\_start=1

/transl\_table=11

/product="Adomet synthetase and MTHF reductase fusion"

/protein\_id="CAC94889.1"

/db\_xref="GI:16197723"

/translation="MSGHFTSESVEGHPDKIADISGTPIDALLREDPSRAVET LITGGVAVADKVTGAYVPIQOLVRDAVLTIGDSAGFDGACGVSISIGSPSD IAGCVDAVESLVEBAAGADIDELDRDAGCGAIFGCATDESPIMLPITLAVL AKRLIDVRNGDLPPTFARTARPGSPEDRAGRATVMDTTRGCAALRPDEVSYTEAG SPDRIVVALTRRIAETTLRPVAHLVAGHSVAELKQIIGRADVGIRDFALRGDPP GDPRGPMSPHEGFEYAHLEVPVRSLSGFTIRMAFPRHPSPMESDIRHVAC RAGADYATTOMEFEGDYLRLDRPAAACCHPIPIEIPMAPVOVRIIRPAEISAVAF TEDLAHRELOAGNPNVGHRIQVGHATVAERLLAAGVGLHCITLNRSTALDIHRL LTRSRSLSVR"

2290..3738

gene /gene="mtmH"

CDS /db\_xref="taxon:41951"

2290..3738

/gene="mtmH"

/function="putative role in an activated methyl cycle"

/codon\_start=1

/transl\_table=11

/product="AdoHcyase"

/protein\_id="CAC94890.1"

/db\_xref="GI:16197724"

/translation="MSADPTDKVADLSAEFGKETTIAHEMPGLMSIRREYAEAO PLACAVTGSILHMYOTAVLITETVALGVOVRAASINISSTOGHAAALVAPGPTD NPRAVPYPAKGETLEBYWCTEQALTWDSPIGGMILIDDGDALILVHKVEYRK DGKVPVDTAENDERRVILDELNRTITDSQKWTQASLSEKRGVETTVHRIYENK RDGVLLPAIVNDVNLVSKEDPNKYGCRSLIDGIRKARDVLLGGKAVVFGDGVK AARSIRGOGARIVTEIDPICALQAMADGVYTTDEVDKADIFITTKNKDIYMA DMARKHOAIVGNIGHNEIDMAGLAKVGYTDEIKQVHTWTFPDGKIIYLSGS RLINGATGHPSPFVMSNFADOTLAQITLPIKPGDVPYGVYTLPKHLDBKVARLHLD ALGVKLTTLRPEQASYIGVDVNDGPKSDHRYR"

BASE COUNT 590 a 1485 c 1197 g 528 t

ORIGIN

Query Match 42.1%; Score 615.2; DB 1; Length 3800;

Best Local Similarity 65.3%; Pred. No. 4.2e-131;

Matches 937; Conservative 0; Mismatches 488; Indels 9; Gaps 2;

QY 34 gagtacagattgacgacatcaaccatccatctgttcctggccgtaaggaacttaccctgct 93

DB 2308 GACTTCAGAGTGCAGCATCTCTCTCTGCGAGTGGCCGACAGGATACCTCTGCC 2367

QY 94 gagaaggaatgacaggtcttactgtctctgtctgtgagcgttactccgtctcaagcattg 153

DB 2368 GAGCAGAGATGCGCGCTGATGTCGATCCGAGGAGATGCGCGAGGCGGACCGCTC 2427

QY 154 aaggtgtcagaatctctgttccctcccaatgacatgcaagaagccgtctcatgag 213

DB 2428 GCGCGCGCGCGTCAACCGCTCTCCGTGACATGACCTGGACACCGCGCTCTCATCGAG 2487

QY 214 acattcaacgtctgtgtctgtatgtatgagtggtcttcgcaacatctctctacacaa 273

DB 2488 ACCCTGTGCGCTGCGCGCGCGCGAGTCCGTGGCTCTCTGACATCTTCTCCACCGAG 2547

QY 274 gatacagccgtgtgtctgtatgtctgtctgtgcccacaagcacaacagaagccagccgt 333

DB 2548 GGCACG 2607

QY 334 atcccagcttgccttggaagggcgaaacatccccaatactgtggaagaaacataccgc 393

DB 2608 GTCCCGGTCTGCGCTGGAAGGCGGAGACGCTGGAGAGTACGTGGTGCACCGAGCAG 2667

QY 394 gctctacatggtcagatgtgc-----aaggcccaacagcaggtgtctgatatgtgt 447

DB 2668 GCGCTACCTGCGCGGAGACGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 2727

QY 448 gatgtacatctctcatctcccaagggcttgaatgaaacagccggtgtgtctcagag 507

DB 2728 GACGCGACCTGTGCTGTCCCAAGGGCGGTGAGTACAGAGAGGAGCGGAGGTCCCTCG 2787

QY 508 ccaacagagctgacacacccgaataccgctgtgtctgtctgtacatcaagcaggtcttc 567

DB 2788 GTTCGACACCGCGCGAGACGACGACCGCGCTACTCTGCAATCTTCAACCGCACATC 2847

QY 568 aaccaagaagaacacactgtgacacacagctgtgtccgagcagatggtgtctcagaag 627

DB 2848 ACCGAGCGGTCCCAAGAGTGGACGACAGTGGCGGCGCGCGAGTCCCGCGGTGACGAGAG 2907

QY 628 acaacacaggtgttccacgctcttaccagctgtgagaagggcgaaactctctctcca 687

DB 2908 ACCACACCGCGCTCCACCGCTGTACAGATGAGGCGCGGCTGTCTGTCTCTCC 2967

QY 688 gccatcaacgtaacagcgtgttaccagaagtcgaatcgaatacatcagctgtgcgc 747

DB 2968 GCGATCAAGCTCAACGACGCGGTACCAAGTCCGAGTTCGACACMAATACGCGTCCGC 3027

QY 748 cactccctatcgatgtaacacgctgtctcgatgtaacgctgtcagcgcgaagaagct 807

DB 3028 CACTCGCTGATCGACGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3087

QY 808 ctgcgtatggtgttaccgagctgtctgcaagggcgctgctcaatccctctcgtggccaagc 867

DB 3088 GGTGTTTGGCTACGCGCGAGCTGCGCAAGG---CTGCGCGAGTCTCGCGCGCAAGGC 3144

QY 868 gctgcgtatcatcaacgaactcgaccacaaactgtgcgtctcaccagctgtcgaatgaaagc 927

DB 3145 GCGCGGTATCTGATCAACCGAGATGACCGCATCTGCGCGCGCGCGCGCGCGCGCGCG 3204

QY 928 taccaggtccgacgcatcgagaagctgtcgaagatgtcgaatctctgttaacatgacaca 987

DB 3205 TACCAAGTACGACCGCTCGAGAGTGTGTCACACAGGCGCGCATCTTATCATCCAGACC 3264

QY 988 ggaactgagatacatctctgtgtgacatgtagccgaatggaagatgagatgtagtgc 1047

DB 3265 GGCACAAAGACATCATCATGCGCGGAGACATGGCGCAGATGACACCAAGCCATCTGTC 3324

QY 1048 gtaacatcgccactcgaataacgaatgatacagatggtcgtcgaataacacagc 1107

DB 3325 GGCACATCG 3384





QY 1112 tccctctgtgagggccgcctcttaacccttgagctgcctcaggtcaaccactctcg 1231

Db 1248 tgcctctgtccgaaggctgcctctgtgaaccttggcaaccgacacccacccatctttg 1307

QY 1232 ttatgtcaatgtcaattacaacaacagacactgcgtcagctgaacctctacgaagaagag 1291

Db 1308 tcatgtccaaactcttttgcgcgcattcagacacattgcccagatcgaactgtttccaaacgaag 1367

QY 1292 gaactctcgaagaagaaggtttacacactctcgaagcatctgatgaagaagtgcctgc 1351

Db 1368 gacagttacgaacacagaggtctacccgtctgccttaagttttccacgaaagagtgacacga 1427

QY 1352 tccacctcgagatcctcgtatgtccacacttaacaagcttaacagaagaagagctgaactaa 1411

Db 1428 tccacagttgaggtctctggcgcgtcagctacacccaactgaccaaaggacagctgagttaca 1487

QY 1412 tcaagctccaagttgaggtctctcaacaagtctgtagtcttaacgttaaa 1461

Db 1488 tcggcgtttacggttcgaagcccatcttaacgacggagacactacccgtactaa 1557

RESULT	6				
AX244105	AX244105	1557 bp	DNA	linear	PAT 28-SEP-2001
LOCUS	Sequence	97 from Patent WO0166573.			
DEFINITION	AX244105				
ACCESSION	AX244105.1	GI:15859169			
VERSION					
KEYWORDS					
SOURCE	Corynebacterium glutamicum.				
ORGANISM	Corynebacterium glutamicum				

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 1557)	Pompey,J.M., Kroeger,B., Schroeder,H., Zelder,O., Haberman,G., Kim,J.W., Lee,H.S. and Hwang,B.J.	Corynebacterium glutamicum genes encoding metabolic pathway

**JOURNAL** Patent: WO0166573-A 97 13-SEP-2001;  
**BASF AKTIENGESELLSCHAFT (DE)**  
**FEATURES** Location/Qualifiers

```

/organism="Corynebacterium glutamicum"
/db_xref="taxon:1718"
101. .1537
CDS

```

BASE COUNT	348 a	440 c	431 g	33
ORIGIN				

Query Match	40.8%	Score 596.4	DB 6	Length 1557
Best Local Similarity	64.5%	Pred. No. 8.5e-127		
Matches 923; Conservative	0	Mismatches 501	Indels 6	Gaps 2

92 ctgaagaagaatgcccagcttattgctctctgtagcgttattccgccttaagcatt 151

Db	174	CAGAGTATGAGATGGCAGGTCTCTCATGCGAATGTTGGCAAGAAATATTCGAGACGAGCAGCTT	233
OY	152	tgaagggtgtcgaagatctctgtgttccctcccatatgacagttccagaagaagccgttccatag	211
Db	234	TGAAAGGGGGCCCGAAATGGTGGTCTTATCCAAATGACGGGTCCAGACGGCCGTGCTTAATG	293
OY	212	agacactcaacagctctgtgtctgtatgtcagaatgagctctcctcctcaacatctctctacac	271
Db	294	AGACCTCTACTCTTTGGGCGCTGAGAGTTGCTTTGGGCTTCTCTCACAATTTTCTCCACCC	353
OY	272	aagatacacagcgcgtgtctgtcatctcgttctgcgcgcccaacaggcacacacagaagaagccag	331
Db	354	AGGATGAGAGCTTCCACGGGCTATCGTTGTGTGGGC---TCCGGCACCGTTCGAAAGACCGCTG	410
OY	332	gtatcccgactctgcctctbgaaggaggaaacactcccaagaatctbgaagaacacatacc	391
Db	411	GGTTTCAGATATTCGGGTGGAAGGTGAGTCACTGGAAGAGTACTGTGTGTGATCAAC	470
OY	392	ggcctctcaatggtgcgaatggtlcaaggcccaagcaggaagttgtcgaatgaatgtgtatg	451
Db	471	AGATCTTGACGTGGGGCGATGAGC---TGCACAAATGATCTCCGACGACGGCGGTGACG	527
OY	452	ctaacctctcatctcccaagggtcttcgaattcgaaacagccggtgtgttccagaagccaa	511
Db	538	CCACCATGGCTTTATTTCCGGTCGGGAAATCCAGACAGCTGGTGTGGTTCCACACAG	587
OY	512	cagaagctfacaacctcgcgaataccgctgtcgttcttctacactcaacgaagtttcaacc	571
Db	588	AGGCCAAGCATTTCCGATGATGATCATGCGAATTTTGGGCATGTCGGTGAATCTTGTCTG	647
OY	572	aagacaagaacacactgtgcaacacagttgtctgcgcgcatgaacggttgttccgaagagacaa	631
Db	648	CAGAGCTTGCGAAGTGGGGCAAGATGCTGAGGCGCTTAAGGGTGTACCGGAGGAAACCA	707
OY	632	caacaggtgtccaacgcctctacacagctcgcgaagaaggagaacactctctcccaagca	691
Db	708	CCACCGGTGTGCACGCCCTGTACCACTCTCCGTTAAGAAGCGGTGCTTCTTCCACAGA	767
OY	692	tcaacgtcaacagcgtctgttacaagttccaagttcgaatacatctacgctgcgcgcact	751
Db	768	TGACGCTCACAGCGCTGTCTCACCACTCCAAAGTTTGTAAACAAGTACGGGACCGCCACT	827
OY	752	cccttatgatgtatcaaacccgtgtctcgcgaatgtaatgacgcgcgcgaagaacagctctcg	811
Db	828	CCCTGATCCAGCGCATTCACCGCGCCACTACATGCTCATGGCGGCGAAGACGTGGTTG	887
OY	812	tcatggtttacgcgcgatgttcgcgaaggtctgcgctcaatccctctcctgtgcgaagcgtc	871
Db	888	TCGTGCGTTACGGCGATGTGTCGCAAGGGCTGCCCTGAGGCTTTTCGACGGCCAGGCGCTC	947
OY	872	gggtatactaacagaagactcgaccacaatctgcgctctcagaagtggtgcgatbgaagctacc	931
Db	948	GGGTCAAGGTCCACGGAAGCTGACCCAAATTCACGCTCTTCACGGCTCTGATGATGGCTACT	1007
OY	932	aggtccgcgcgcacatcgaggaagtgtcaagaatgtlcgatatacttcgttatactgcacagaa	991
Db	1008	CTGTGTGTAACCGTTGATGAGGCGATCGAGGACCGCCGACATCGTATGATCACCGCGACCGGA	1067
OY	992	actgtgatatcaatctctgttgaacatgataggcccaagatggaaggaataagctatgttgta	1051
Db	1068	ACAAAGACATCATTTCTTCGACACAGATGCTCAAGATGAAGAGATCACGCTCTGCTGGGCA	1127
OY	1052	acatgccaacttcgaataacgaatatgatacagatgagctctcatatgaatatccacagacatca	1111
Db	1128	ACATGGGTCACTTTGATTAATGAGATGATGATATTCGATTCCTGTTGACACGGGACGACGTC	1187
OY	1112	agcacatcccaatcaagccgaatcaagacatctgtggaatctcccaagatgtgcacgctatcc	1171
Db	1188	CCGCGACCAAGATCAAGCCACAGGTGACGACATTTCTTCACCGCGTGTGCTCATCA	1247
OY	1172	tcctcttctgtgaggcgccctcttaactgtgtcgtctaaaggtgcaccatcttgcg	1231





RESULT	8	AXI20920	AXI20920	1422 bp	DNA	Linear	PAT 11-MAY-2001
LOCUS		AXI20920					
DEFINITION		Sequence 836 from Patent EP1108790.					
ACCESSION		AXI20920					
VERSION		AXI20920.1					
KEYWORDS		GI:14037635					
SOURCE							
ORGANISM		Corynebacterium glutamicum.					
REFERENCE		Corynebacterium glutamicum					
AUTHORS		Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;					
		Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;					
		Corynebacterium.					
TITLE		1 (bases 1 to 1422)					
JOURNAL		Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,					
		Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.					
		Novel polynucleotides					
		Patent: EP 1108790-A 836 20-JUN-2001;					
		KIOMA HAKO KOGYO CO., LTD. (JP)					
FEATURES		Location/Qualifiers					
source		1..1422					
BASE COUNT		/organism="Corynebacterium glutamicum"					
ORIGIN		/db_xref="taxon:1718"					
		309 a 414 c 401 g 298 t					
Query Match		40.7%; Score 594; DB 6; Length 1422;					
Best Local Similarity		64.5%; Pred. No. 3e-126;					
Matches 920; Conservative		0; Mismatches 500; Indels 6; Gaps 2;					
QY	32	tcgagtaacaagattgcccagatcaactccatcagttctgcgcgtaaggaaattacacctg	91				
DB	2	tggactttcagaagtgcgcgattttacatcagacagagcagagctaccagattcgctttg	61				
QY	92	ctcgaagaagaaatccagaagctcttatggtcttcttcgtctgaaggttatctcgccttcacgcat	151				
DB	62	cagagctatcagattgccagagctctcattccatttggcgaaagaaatttcgacagacagaccctt	121				
QY	152	tgaaggtgtcagaatcttgltccctccacaatgacagltccagacagccgctctcatcg	211				
DB	122	tgaagggcccccgaattgctggtttcattacatgacgggtccagacggccgcttcattatgg	181				
QY	212	aagaccacacactcttgctgcgatagtcaagltgggcttccttcgcaacatctctctaacac	271				
DB	182	aagaccttactgcttttggcgctgaagttcgtttggccttcttcacacattttctccacc	241				
QY	272	aagatacagcgcgtgctgtctatcgttcttcgcccacaagcagacacacagaagaacgacg	331				
DB	242	agagatgagctgcagcggcgttatcggtgtgggc---tccggaccctgcgaagaccagctg	298				
QY	332	gtatccagctcttcgctctgtaaggggaaacactcccaataatcttggaacacatacc	391				
DB	299	gtgttcacagatttcgctgtaaggggtacgactgagagagatctgctgtgcataaac	358				
QY	392	gcgcctccacatggtcagatggtcgaagggccacagagaggtgttcgaatgtgtgtgatg	451				
DB	359	agattttcagctggggcgatgagct---gcgaataatgatcttcgacgacggcggtgacg	415				
QY	452	ctacactccctatctccaagggcttcgaattcgaaacagccggtgtgtctcagagccaa	511				
DB	416	ccaccatggctgtttattccgcggcggaattacagagacagctggttcggtccacgacag	475				
QY	512	cagaagctacaaacctcgatatccgtgtgcttcttcacactcaagcagagcttccaacc	571				
DB	476	agggccaacgattccgattgagatgcattcttggggcattgcctggaaggttcttgcgtg	535				
QY	572	aagaagaagaacccattgacacagttgcttcgcgacatgaacggtgttccgaaagagcaa	631				
DB	536	cagagcctgcgaagtggggcgaagatgcgtgagggccgttaagggtgttcacagaggaacca	595				
QY	632	caacaggtgtccacgcgctctacagctcgtgagaagaaggcaactcctctccagcca	691				

D	596	CCACCGGCTGCACCGCCTGTACCACCTTCGGTGAAGAAGGGTGCTGCTGCCATTCCACGCA	655
Q	692	tcaacgttcaaacgacttgttacaagaatccaaagttcagataacatactcagctgtccgcact	751
D	656	TGAACGTCAACGACGCGTCCACCAAGTCCAAGTTTGATTAACAAGTACGGCACCCGCCACT	715
Q	752	caccttcagtcatgaataaccgttgtctccgatgtccagatcgaggcggaagacacactctcg	811
D	716	CCCTGATGAGCGCATCAACCGCGCACATGCAATGCTCATGTGGCGGCAAGAACGTGCTTG	775
Q	812	tcaatgggttiacggcgaatgtcggcaaggctgtcgctccaattcccttcgtggccaaggctc	871
D	776	TCTGGGGTTACGGGATGTGCGCAAGGGCTCGCTGAGGGCTTTGACGGCGACGGCCCTTC	835
Q	872	gcgttatcatcacagaactcgaaccgaatctcgctctccaggtccgacatggaaagctacc	931
D	836	GCGTCAAGGTCAACGAACTCAACCAATCAACGGCTCTTCAAGCTCTGATGGATGGCTACT	895
Q	932	aggtccgcgcacatcgagaagctgtcgaaggaatgtcgaatatctcgttiacatgycacaggaa	991
D	896	CTGTGGTCAACGGTGTGATGAGGCCATTCGAGGACGGCGCAATCGTATCAACCCGACGGCA	955
Q	992	acttcgatatcatctctgttgaacatgataggccccagatgaagaataaagctatttgcgta	1051
D	956	ACAAGGACATATATTTCTTTCGAGCAGATGCTCAAGATTAAGATCAACCTCTGCTGGGCA	1015
Q	1052	acatcggccacttcgatacaagaatttgaataagatgacctatgaatatccaaagcatca	1111
D	1016	AACATCGGCTACTTTGATTAATGAGATCGAATGCAATGCCCTGTGGACCGCGACGACGHCA	1075
Q	1112	agcacatcccacataaacgccagaatacagcatatgttggaaattcccagatbgccaagctatcc	1171
D	1076	CCCCGACACGACATCAAGGCCAAGGTGCAGAGATTCACCTTCCACCGGTGCTCCATCA	1135
Q	1172	tccctctgtgcgaaggccgcctcttcaacctgtgcgtcgtcacaagytcaaccatcttcg	1231
D	1136	TGCTGCTCTCCGAAGAGTGGCCTGTTGAACCTTTGGCAAGCCACGCAACCATCATTTTG	1195
Q	1232	ttaatgcaatgbcataccaacaacagacactcgcctaagctcgaccttagaanaaagag	1291
D	1196	TCAATGCCAACATCTTTCCCCTCATGACCAATTTGCCCAATGCAATGTTTCCAAAACGAAG	1255
Q	1292	gaaatctcgagaagaagtttacacacttcgaaagcatctcgatgaagaagtgcgtgcc	1351
D	1256	GACACTACGAGAACGAGGCTTACGCTCTGCTTAAGGTTCTGACCAAAAAGGTGGACGCA	1315
Q	1352	tccaaactcgaatctctcgaatgtccaaacttaaaaagcttacaaagaagcagtgaacta	1411
D	1316	TCCAGGTTGAGGGCTCTCGCGGGTCAAGCTCACCGAAGCTACCAAGGACGAGCGTGAATA	1375
Q	1412	tcaaagttccagttbaaggttcttcaacagatcgtatgacttaccogtta	1457
D	1376	TGCGGCTTAGACCTTGCAAGCCCATTCAAAGCCGAGCACTACCGCTTA	1421
RESULT	9		
AF262755		1488 bp DNA linear BCT 30-MAY-2000	
LOCUS	AF262755		
DEFINITION	Mycobacterium bovis S-adenosyl-L-homocysteine hydrolase gene,		
	complete cds.		
ACCESSION	AF262755		
VERSION	AF262755.1 GI:8101959		
KEYWORDS			
SOURCE			
ORGANISM	Mycobacterium bovis.		
	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
	Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;		
	Mycobacterium; Mycobacterium tuberculosis complex.		
REFERENCE	1 (bases 1 to 1488)		
AUTHORS	Pavari,S.N. and Nayak,R.		
TITLE	Complete genomic DNA sequence of the S-adenosyl-L-homocysteine		
	hydrolase gene from Mycobacterium bovis BCG		

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1488)  
AUTHORS Pawar, S.N. and Nayak, R.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAY-2000) Microbiology and Cell Biology, Indian  
Institute of Science, Malleshwaram, Bangalore, Karnataka 560012,  
India

## FEATURES

## source

## CDS

Location/Qualifiers  
1..1488  
/organism="Mycobacterium bovis"  
/strain="BCG-Pasteur"  
/db\_xref="taxon:1765"  
1..1488  
/codon\_start=1  
/transl\_table=11  
/product="S-adenosyl-L-homocysteine hydrolase"  
/protein\_id="AAF72670.1"  
/db\_xref="GI:8101960"  
/translation="MTGNLVTNKSLTPDVRNGIDFKIADLSTADFGKRIERIEHEMP  
GLMSIRREAYOPLKGRIRISSLSIMTVOTAVILFLLDGAENVKMSIESTODHA  
AAAVVGHGTPDEPKVPEVFAWKGETLEEWAAQMLTWPDPKANNILDDGDA  
TMLVLRGMYEKAGVVPABEDDPAEWKVFLMLRTREFDKDKWKIASVGVTEE  
TTLGVLRLOYFAAGDLAPAINVDSVTKSKEDNKGTRHSLIDGINRTDLIGK  
KVLICGVDVGKCGEAKKGOGARVSTEIDPILNLAAMEGSDVYVEAIGDADIV  
VTATGNKDIIMLEHKKAKMDHAIIGNIGFONEIDMAGIERSATVYVNPQDUMTF  
GPTGRSTIVLESRLINLGNMATHGHPSPVANSFPANOTIAOIELMTYNDEYTRLP  
KHDEKARIVEALAGHLLTKTKSOAEYLGVDEGPKPDHYR"

BASE COUNT 313 a 475 c 474 g 226 t

ORIGIN

Query Match 36.6%; Score 535.4; DB 1; Length 1488;  
Best Local Similarity 61.3%; Pred. No. 9,2e-113;  
Matches 879; Conservative 0; Mismatches 551; Indels 3; Gaps 1;

QY 32 tcgagatcagatttcgcagatcaactccatcgtctcgtcgcgttaaggaacttaaccttg 91  
DB 56 TCGACTTAAATGCGCGACGACTGTCACTAGCGGATTTCGCGCGCAAGAACTCGGATCG 115  
QY 92 ctgagaagaaatgcaggtcttatgtctctcgttgtagcggtatctccgtcttaagccat 151  
DB 116 CCGAGCAGGATGCGCGCGCTGATGCTCGCGCGCGAGTATGCCAGAGTGCAACCCC 175  
QY 152 tgaaggtgtcagaatctcgtctcgtctccatcagtcagtcagtcagtcagtcagtcagtc 211  
DB 176 TGAAGGAGGCGCGGATGCGGTGCTGATGCAATGCGGTGCAACCGCGCTGTGTTCG 235  
QY 212 agaacatcacagctcttggtgtgtgtatgtatgtatgtatgtatgtatgtatgtatgtat 271  
DB 236 AAGCCCTACCGCGCTGCG 295  
QY 272 aagataagccgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 331  
DB 296 AGGATCACG 355  
QY 332 gttacccagctctcgtcgtgaagagcgaaacatcccaagaatctgtggaacacataacc 391  
DB 356 GTGTCTCGGTGTGCGGTGAAGGCGCAACGCTCGAAGAGATCTGTGTGCGCGCGCG 415  
QY 392 ggcgtctcacatggccaaga----tggtaagagcccaagcaggtgttcagatgtgtgtg 448  
DB 416 AGATGCTCACCTGGCGGACCCCGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475  
QY 449 atgtactactcctcatctcagaaggtctcgaatctcgaacacagcggtgtgtctcagaagc 508  
DB 476 ACGGCACCATGTTGTGTGCGCGCATGCAATGAGAAAGCGCGCGCGCGCGCGCG 535  
QY 509 caacagaagctcgaacactcgaatccgctgctgtctgtctgtctgtctgtctgtctgtct 568  
DB 536 CCGAGGAGGACGACCCCGCGAGTGAAGGTTCCTCTGCACTGCTACGACCGCGCTTCG 595  
QY 569 accaagaacaagaacactgtgcacagttgtctccggtcagtcagtcagtcagtcagtcagtc 628

DB 596 AGACCGCAAGAGCACTGACCAAGATAGCCGAGTCCGTCAGAGGCGCTCACCGAGAGA 655  
QY 629 caacacaggtgtccacgcgcctcaccagctcgaagaagagggcaactccctccaccg 688  
DB 656 CCACCAACCGCGGTGCTGCGGTCTACCAATTCGCCGCGCGCGCGCGCGCGCGCGCG 715  
QY 689 ccaataagctcaacagcagctgttacaagagtcgaatcgaatcgaatcgaatcgaatcgaat 748  
DB 716 CGATCAACGTCACAGCATCGGTGACCAAGTCCAAATTCGACCAAGTACGCGACCTCG 775  
QY 749 acccttcattcgtatgtatcaacagctgtcttcagatgtatgtatgtatgtatgtatgtat 808  
DB 776 ACTCCCTGATCGACGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 835  
QY 809 tctcagtggttaccagcagctgtcgaagagctcgcgcctcgaatccctccgttgcgaagcgc 868  
DB 836 TCATCTGCGCGCTACCGCGAGCTCGGTAAAGGCTGTGCGGAGCGCATGAAGGCGCG 895  
QY 869 ctgcggttaccatcaacgaactcgaaccatctgcgtctccagctgtccatgtgaagcgc 928  
DB 896 CCGCGGCTCTCCGTCACCGAGATGACCGGATCAACGCGCTGCGAGGCGCATGATGAGG 955  
QY 929 accaggtccgcagctcgaagaagtcgtcaagatgtcgtatctgttataatcgaagc 988  
DB 956 TCAGCTGTGTCACGCTGAGAGAGGCGCATCGGCGCGCGCGCGCATGCTGTACCGCG 1015  
QY 989 gaacatgcatacatctcgttgcataatgtatgcagtcgaagtcgaagtcgaagtcgaagtc 1048  
DB 1016 GCACAAAGCATCATCATCTGACGACCATTAAGCGCATGAGGACCGCATGCTCTG 1075  
QY 1049 gtaacatgcgcacatctcgaataagaaatgtatcagatgtcgtcgaataagtcgaagtc 1108  
DB 1076 GAATATCGCGCGCATCTGACAGCAGATGCGCATGCGCGCGCGCGCGCGCGCGCG 1135  
QY 1109 tcaagcattcccaatcaagcagaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1168  
DB 1136 CACGGGTCAACGTCACAGCTTCAAGTCACTGTGACCTTTGGCGACACGCGCGCTG 1195  
QY 1169 tctcctcttctgtcgaagagcgctctcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1228  
DB 1196 TCATCTGCTGTCGAGGCGCGGTGCTGAAACCGTGGGCAACCGCGCGCGCGCG 1255  
QY 1229 tctgtatgtcatatgtatcacaacacgaacactcgtcagctcgaatcgaatcgaatcgaat 1288  
DB 1256 TCGTGTATGAGCAACAGCTTGTGTAACGACGATGCGCGCATGAGCTGTGACCAAGA 1315  
QY 1289 gaggaaatctcgaagaagaggtttacacactcgcgaagcactcgaatcgaatcgaatcgaat 1348  
DB 1316 ACAGCAGTACGACCAAGAGGTGTACCGGCTGCTCCCAACCTCGACGAAAGGTGCG 1375  
QY 1349 gctccacactcgaatctcgtatgtccacttacaagcttacaagcagagcaggtgtact 1408  
DB 1376 GAATCATGTGAGGCGCTTGGCGGTGACCTGACCAACGCTACCAAGAGGACGCGCGAAT 1435  
QY 1409 acataaggttccagttcgaatgaggttcctacaaagtcgtatgtaccgttataa 1461  
DB 1436 ACTCTGCGCTGACGCTGAGAGGTGCTTCAACAGCGGACCACTACCGGTACTGTA 1488

## RESULT 10

AX023852 1488 bp DNA linear PAT 15-SEP-2000  
LOCUS  
DEFINITION Sequence 23 from Patent WO0021983.  
ACCESSION AX023852  
VERSION  
KEYWORDS AX023852.1 GI:10184196

## SOURCE

ORGANISM  
Mycobacterium tuberculosis.  
Mycobacterium tuberculosis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE  
1 (bases 1 to 1488)  
AUTHORS Skjot, R.L., Andersen, P., Rosenkrands, I., Weidinger, K., Okkels, L.M.,



gene /gene="MT3342"  
/note="identified by glimmer2; putative"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="AAK47684.1"  
/db\_xref="GI:13883162"  
/translation="MERMLRLTILFLGAVILGACSAVSPASQAPQITVERPVPNL  
KPSGMDVDYLREFIKATADPARNRLAARQPTTESASNMDDAGSLIDHVFVE  
RSAAKVSYTMADILIGSDVGVEFAEQQLDPGPIELVKTIDGWMIDRLPNKVFLE  
WQOQETIKRNTLTFTADPTGKTIVDPDPVAVASDMQRLATELVSKLAGRPEAKRY  
NLLAPLRLRGPVVRADGKSGIGIGYSGAGVADKRLSTDPHSROLAAQIITWTLA  
QRIYVPVFIINADGAPLEDFEAFEGWTTSDVAATDPVADGAAGLAVIGSLVAMPA  
QAVTPAFGFMPEPQOTAAVRSRGROVAVYTLIGAGDPDEAGLWGDIGGEAVQSA  
DGLSLSRWSGLDDAVVWVYDNTVYLRALODPASGQPARPUDSTAVASRFPQAIN  
QLSPDGRPAANVTIGQVILAGYEQTQAGFPALTYPRLGEGLSSVSLSWRTDDIV  
VTRIDAHPSAVSYVINDGVNSDAPSRGLQPTALIAANPSTVYVAGPGVILMTSASV  
RPGADVPLGLWPGCAVPLPG"  
/gene="MT3343"  
/complement(4370. .6073)  
/note="MT3343"  
/gene="MT3343"  
/note="similar to PID:556614 PID:2072712 GB:AL123456;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="sensor histidine kinase MtrB"  
/protein\_id="AAK47685.1"  
/db\_xref="GI:13883163"  
/translation="MIFSSRRIRIGRRGRSGPMTGRLSALSRAAVAMRRSLQLRYVA  
LTLGLSLAVLILGLFVLTQSQVYNNRVLDIRAIDQIERARTVSGIVNGEETSLDS  
SGLQARNTVSKTDPASGAGLADPAVLMVGDGGRASAPVQPVNALGKFAVA  
SOAQAQATVQTEGSGPALLIGTFLTSVNLDELITPLASQOATITLVRGFMATG  
GLVTLVGLIALILVSRQVYVRSASRAEFPAEHLSEPMVREDMDRAIYASPN  
DMESLRLQALDLEEGNLORRFTSGRSTLEPTPLTYVMADLTIDHSDIDPTLR  
STELMSELDRFETLNDLIELSRDAGYAEISVEAVDRTTYNNALGVNGHIAEEAG  
ITELLVDPAEYIAEVDARKVERILRLNLAINAIDHAENKPVRLIMADEPTVAVTAD  
YGLGREGKEELVSRFMRSDPSRVRSGGTGLGLAIVSDALDHQRLTEANEPGEG  
ACFRLTLPLVNGHKVYTSPLPMKPIPOVPLQVADPNQPMPEYKERQRPREHAES  
G"  
/complement(6123. .6830)  
/gene="MT3344"  
/complement(6123. .6830)  
/gene="MT3344"  
/note="similar to GB:U01971 PID:509815 PID:2072713  
GB:AL123456; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="DNA-binding response regulator MtrA"  
/protein\_id="AAK47686.1"  
/db\_xref="GI:13883164"  
/translation="MSRGGDTMDYMKRILVYDDASIAEMLTIVLRGEGFDYVIGD  
GKALVAREIRPDLVLDLMLPGNMGIDVCRVLRADSGVPIYMLTAKPTDYVYGL  
ESGADYIMKPFKEKELVAVRARLRNRDDPEAKMLSIADVEIPAHKTRNGEOL  
LTFLEPLVLAIAKKRPROVFTRDVLLBEQWKGIRHADRRLVAVHQRLLAKVEKDE  
PNVYLVIRGVGYAKRGP"  
/complement(6879. .7523)  
/gene="MT3345"  
/complement(6879. .7523)  
/gene="MT3345"  
/note="similar to GB:AEO00782; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="thymidylate kinase, putative"  
/protein\_id="AAK47687.1"  
/db\_xref="GI:13883165"  
/translation="MLTAIEGVDAAGKRTIENKLSGARFAAGRSVATLAFPPYGQSA  
ADIAARHENAAGCAAAAYQRIETARLGLPKPDQOVLLVASIELAGERSGRQRPDGR  
ARNRIRDEDELQRTGCAVYAEVLAAGWGWGRWLTVGADVDPGLATATLAPPDVPS"  
/complement(7620. .9107)  
/gene="MT3346"  
/gene="MT3346"



LOCUS MTCY20B11 36330 bp DNA linear BCT 03-AUG-2001  
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 139/162.  
ACCESSION Z95121 AL123456  
VERSION Z95121.1 GI:3261742  
KEYWORDS  
SOURCE  
ORGANISM Mycobacterium tuberculosis H37Rv.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;  
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;  
Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE  
AUTHORS Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skellon, S., Squires, S., Squires, R., Sultston, J.E., Taylor, K., Whitehead, S., and Barrell, B.G.  
TITLE Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence  
JOURNAL Nature, 393 (6685), 537-544 (1998)  
MEDLINE 98295987  
REFERENCE 2 (bases 1 to 36330)  
AUTHORS Parkhill, J.  
TITLE Direct Submission  
JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk  
COMMENT On Jun 27, 1998 this sequence version replaced gi:2072692.  
Notes:  
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.  
(URL: <http://www.sanger.ac.uk/Projects/M.tuberculosis/>) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.  
Implemented in Tblastx (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.  
CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.  
FEATURES  
source  
1. 36330  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
<1. 34116  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
/db\_xref="taxon:83332"  
/clone="Y20B11"  
complement(90..848)  
/gene="RV3226c"  
complement(90..848)  
/gene="RV3226c"  
/note="RV3226c" (MTCY20B11.01c) len: 252 aa, unknown, but similar to various hypothetical proteins, e.g. BSB0011\_212 Bacillus subtilis O31916 YOWW PROTEIN (224 aa)  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein RV3226c"  
/protein\_id="CAB08327.1"  
/db\_xref="GI:2072693"  
/db\_xref="SPTREMBL:O05872"

misc\_feature  
1914..1970  
/gene="aroA"  
/note="PS00885 EPP synthase signature 2"  
2252..3244  
/gene="RV3228"  
2252..3244  
/gene="RV3228"  
/note="RV3228" (MTCY20B11.03) len: 330. Function: unknown but similar to several hypothetical bacterial proteins, Yj60. FASTA best: YJ60\_HAEIN P45339 Hypothetical protein h11714 (346 aa) opt: 355 z-score: 404.9 E(): 1.5e-15; (31.7% identity in 281 aa overlap); contains PS00017 ATP/GTP-binding site motif A"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein RV3228"  
/protein\_id="CAB08329.1"  
/db\_xref="GI:2072695"  
/db\_xref="SPTREMBL:O05873"  
/translation="MRPGVDSDVYKRSGRSSPRKTRPEHADAAMVSVNDRG KGCYIGRPRDRTTAMRARELRPIYGVGDVYVGLSGPDTLARIYKRRPRITVL RRTADDTPTERRVYANADLLIVALLADPPRIVGLDRAIIAIVAGGLTTLTKT DLAPAEFGKQFQDELVTVAAGVADLLVADILAGKITVLLGHSVGKSLVRLVY PADRAVGEVTEITLGRGHTSTRSVALPLGDTLSSGWVIDPGIRSFGLAHIQPNVL IAFSDLAETATRECPRGCHMGPPADPECALDITSGPAARRAAARRLAVLSQT"  
2861..2884  
/gene="RV3228"  
/note="PS00017 ATP/GTP-binding site motif A"  
complement(3277..4560)  
/gene="desA3"  
complement(3277..4560)  
/gene="desA3"  
/note="RV3229c" (MTCY20B11.04c) len: 427. Function: desA3, but shows some similarity to desaturases. FASTA best: D90914\_112 Synecocystis 008671 LINOLEYL-COA DESATURASE (EC 1.14.99.25) (359 aa) opt: 319 z-score: 380.8 E(): 3.2e-14; (25.1% identity in 295 aa overlap).  
/codon\_start=1  
/transl\_table=11  
/product="desA3"



Query Match	Best Local Similarity	Matches	879; Conservative	36.6%; Score 535.4; DB 1; Length 36330; Pred. No. 1.1e-112; Mismatches 551; Indels 3; Gaps 1;
DB 27118	TCGACTTTAAAGATCCGACACCTGTCATCAGCGGATTTGGCCGCCGCAAAAGAACTCCGGATGCG	27059		
DB 27058	CCGACACAGATGCGCCGCTGATGCTGCGCTCGGCGCGAGTATGCGCGAGTGCACACCC	26999		
DB 26998	TGAAGGGGGCCCGGATCTCGGGTTGCGCTGCACATGACGCGTGCAGACCCGGTGTGATGAC	26939		
DB 212	agacacacagctcttgctgatgtcagatggtgcttcctcgtgacacatcctctctacac	271		

Dh	26938	AAACCCATACCCGCTGGGCGCCCAAGTCCGCTGGGCTCTGTGCAACATCTTTCACACC	26879
Qy	272	aaagtaacgcgcgtgctgtctatctgttttcgcggcccaacagcgacacccagaagacgcagccg	331
Dh	26878	AGGATCACGCGCGCGCCGCTGTGTGTGGCCCGCACGGCACACCCCGACGACGCCAAG	26819
Qy	332	gtatcccgattctcgctctggaagggcgnaaacctcccgaatctatggyagacacatacc	391
Dh	26818	gTmTCCGGTGTGCGGTGAAGGGCGAGACGCTTGAAGATAGTGGTGGGCCCGCGAGC	26759
Qy	392	gcgccttcaacttgcgcaca----tgtcaagcccaacagcaggttgtcattatgttgtg	448
Dh	26758	AGATGCTACCTGGCGGGAACCCCGACAAGCCGGCCAAATATCTTCATACGGCGGTG	26699
Qy	449	atgtacactctctcaatccaaaggtctcgaaattcgaaacacgcggtgtgtttccagagc	508
Dh	26698	AGCGCACCATGTGGTGTGGCGGCGCATGATGAGAAAGCGCGCGGTGGTGCGCCGCG	26639
Qy	509	caacagaagctgtgcaactcgaatacgcgtgcgtctgtctgtacactcaagcaggtcttca	568
Dh	26638	CCGAGAGAGAGAGACCCCGCGAGTGGAGGTCCTTCTCAACCTGTGTACGAGACCCGCTTCG	26579
Qy	569	accagaagaaaacacactcggcacaacagttgtctgcgcgcatgaacggtgttccagaagga	628
Dh	26578	AGACCGACAAGGACAAAGTGGACCAAGATAGCCGATCGGTCAAGGGCCTCACCGAGAGA	26519
Qy	629	caacacacaggtgtccaccgcctctcaaccgctcgaagaagggcaactcctctccag	688
Dh	26518	CCACACCGGGGTGTGGGCTCTCAATTCCGCGCGCGGGAGATGTGGCTTCCGCG	26459
Qy	689	ccataaagtaacagcagcgtgttaaaagttccaaagtgtgataaacgtcgaatgcgcgc	748
Dh	26458	CGATCAAGGTCAACAGCTGGGACCAAGTCCAAATTGACACAAAGTACGGACTCGGC	26399
Qy	749	actccctcatcgatgtgataccaacggtgtctccgaatgtatgatacgcgcgcaagcagctc	808
Dh	26398	ACTCCCTGATGACGCGCATCAACCCGCGACCGACCGCTGTATCGCGGTAAAGAGTCC	26339
Qy	809	tcgtcatggttaacggcgatgttcgcaaggggtgtgcctcaatcctctccgttgcgaagcg	868
Dh	26338	TCATGTGGGGCTACCGGCAACGTCGTAAAGGGTGTGGCGAGGCGATGAGAGGCCAGGGAG	26279
Qy	869	ctcgcgtatacatacagaacatcgcgaaccacatctgtgcctccaaaggtctccatgaaagct	928
Dh	26278	CGCGGGTCTCGGTCAACCGAGATCGAACCGGATCAAGCGCGTGTGAGGCAATGATGAGGGCT	26219
Qy	929	accaggtccgcgcgcatcgaagaagtcgctcaagaagatgtgatatacttcgttatactgacag	988
Dh	26218	TCGAGCTGGTCAACCTCGAGGAGGCCATTCGGGGAGCCGACATCTGTAAACCGCGACCG	26159
Qy	989	gaacatgcgatatcatctctgtctgatactgatacgcgaatgaagaataagactattgtcg	1048
Dh	26158	GCATCAAAAGACATCATATGCTCGAGCACATTTAAAGCGATTAAGAGAACACACGGGATCTGG	26099
Qy	1049	gtaacatcggccacttcgataacgaanaattgatatacagaatgacctgaataccccagga	1108
Dh	26098	GAATAATCGCGCACTTCGACAACGAGATGACATGCGCGGTGAGCGCTCTCGGGCGCA	26039
Qy	1109	tcaagcacatcccaatcaagaacgaataagacatcttggaattcccaagaatgcacagcta	1168
Dh	26038	CACGGGTACAGTCAAGCTCAAGGTTCGACTGTGACCTTTGGCGACACGGGCCGCTCGA	25979
Qy	1169	tctcctcttctgtagaagcgcgcttcttaaccttggtcgtctacaggtlcaaccatctt	1228
Dh	25978	TCATGTGTGTGTCCGAGGGCGGCTGTCTGAACCTGGGCAATATCCACCGGGCACCCCTGT	25919
Qy	1229	tcgttatgtcaatgtcatctcaacaacacgaactcgctagtctgtaacctctacgaaga	1288
Dh	25918	TCGTGTTGAGCAACAGCTTCGTAAACAAAGCATGCCCAAGATCGAGCTGTGACCAAA	25859
Qy	1289	gaagaaatctagaagaaggtttatacacacttccgaagatctcgaatgaagaattgcgc	1348
Dh	25858	ACGACGAGTACGACAACGAGGTGTACCGGCTGCCAAGACCTCGACGAGAAAGTGGCT	25799







TITLE	SOURCE
The influence of aluminum on histone, heat shock and S-adenosyl-L-homocysteine hydrolase gene expression in tolerant and sensitive cultivars of wheat Unpublished (1993)	Location/Qualifiers 1. .1708

5' UTR	gene	mRNA	CDS
1. .1708	/gene="Sh6.2"	1. .1708	/gene="Sh6.2"
<1. .1708	/gene="Sh6.2"	<1. .1708	/gene="Sh6.2"
19. .1476	/gene="Sh6.2"	19. .1476	/gene="Sh6.2"
/codon_start=1	/product="S-adenosyl-L-homocysteine hydrolase"	/protein_id="AA34303.1"	/translation="MALSVKTSRGREYKVDLFDQDFGRLELEAVEPMGLMCRPT EFGPQPFKRGARISGSLHMTIQFVLLLETLALGEVRWCSNIFSSDHAALAIAD SAAVPAAMKEETLEEMWCTERCIDMDGGGPDLYDDGDADLLIHGVAKEEFERS KGVPESTDNDPEKIVLTITRDSIKDASHYRKKMERLVYSEETTYGKRLYOMEG SCTLFPAINVNDVYTKSRDNLGCRSLPDDGLMRADVMATGKAAVYCGYDVGKG CAALKQAGARVLTETDPLCALQALWEGIDILTEEDVSEADIFETTGKNIIMVD HRRKKNNNAIVCNIIGHFDEINIMNGLETTPYKRTTIKQTDWPEPRTGIIIVAE GLIMNIGCATGHPSPVWSCFTFNQVIAQLELMEKASGKYEKKVYVLPRLHDEKVALM HGKIGARLTKLKSSDYSIPIEGPYKLRURY"
3' UTR	/gene="Sh6.2"	1708	1477. .1708
polyA_site	/gene="Sh6.2"	1708	
BASE COUNT	344 a 532 c 474 g 358 t		
ORIGIN			
Query Match	36.2%; Score 529.6; DB 8; Length 1708;		
Best Local Similarity	62.9%; Pred. No. 2e-111;		
Matches 911; Conservative	0; Mismatches 489; Indels 48; Gaps 4;		
34	gagtaagaattcgcgacatcaacctccatgttctcgcgcgtaagaacttaccctgtc	93	
55	GAGTACAAGTCAAGGACCTCTTCCAGGCGCACATTGCGGCTCGAGCTCGAGCTGCC	114	
94	gagaaagaaatgcaggtcttatgttctcttcgtgaagcgtltaatcgccttcaagcatg	153	
115	GAGGTGCAATGCCCGGCTCATATGCGCTGCGCCGACCGAGTTCGGCCCTTCGACCCCTTC	174	
154	aaggtgtgaagaatctctgttctccctcaactgaagtcagacagacagcgttccatcagag	213	
175	AAGGGGCGCCGGATGTCGCGCTCCCTCCACATGACCATTCACACCGCCGCTTCATCGAG	234	
214	acaactacaagctcttggctgtagtctcagaagagcgttctcctgaacaactctctacaaca	273	
235	ACCTTCACCGCCTCGGCGCGGAGGTCCGATGCTCTCTCTGCAACATCTTCTCCAGCGAG	294	
274	gataaagcgcgtgtctatcgttgtgtgtcggtgccaacagacacacagagaagcagccgt	333	
295	GACCAAGCGCGCGCCGACATCGCCCGGACATCCGCGG-----	331	
334	atccgaagcttcgcctcggaagagcgaaacactccagaatacttggaagaacataaccgc	393	
332	-----CCGTTTCGCTTGGAAGGCGGAGACCTTCGAGGAGTACTGtGTGACACGAGCGC	387	
394	gctctacaatgagcagaatggtlcaagagcccaacagcaggttgttcgatgatgtgtgtagct	453	
388	TGCTTCGATCGGGGCGTGGCGGCGGCGCCGACCTCATTCGTGACAGACGGCGGTGAGCGC	447	
454	acactcttatctcaagaagctc-----gaattcgaaacagcgcgtgtgtcgtc	501	

D	448	ACGTCGTCATCCACGAGGGCGTCACAGGGCCGACGAGGAGTAGTTGCAAGAAATCCGGCAAGTT	507
Q	502	ccagagccaacaagaagcttgaacaactcgataacccgtgcttcttgctacataacgaag	561
D	508	CCCGACCCGGAGTCCACCCACCAACCCCGGAGTTTCMAGATCTCCTCCACATCATCCGGAC	567
Q	562	gtcttcaacccaagaacaagaaccactgacacagcttgcctgcgcatgaacggtgttcc	621
D	568	GGGCTTCMAGACCGACGCGCCAGCAGATACCCGAMAGATGAGGAGAGGCTCGTCGGGTGTTCC	627
Q	622	gaagagacacaacaagtgltccacgcgccttaccagctcgaagaagagaggaacactctc	681
D	628	GAGGAGACACACACCGGCGCTCAAGAGAGCTTACACAGATGACGAGAGTCGGCACCCCTC	687
Q	682	ttccccaagcataaagtcagaagcgtgttacaagtcacaagtlcgaataacatacagcc	741
D	688	TTCCCCCGCATACGTCACAGCATCTCCGTACCCAMAGCAAGATTGTGCAACCTTTAAGGT	747
Q	742	tgccgcacactccctatacgatgatacaacgcgtcttcgaatgataatgcgscgaag	801
D	748	TGCCGTCACTCGCTCCCTTATGTTGTCTTATGAGGGCCACTGATGTTATGATCGCCGAG	807
Q	802	acagctctgcatacgaagtacgcggaatgcgcgaagggctgcgcatacctccctcgtgc	861
D	808	GTCCGCGTGTCGTGGCGGTACGGTGAATGTGGCAAGGGCTGTCCCGCGCATCAACGAC	867
Q	862	caagcgctcgcggtatatacacaagaacctcgaccacatcgcgctctccagctctgcac	921
D	868	GCTGTGCCCCGTTGATCTGTGACAGAGATGACCCCATCTGTCCTCCCTTCAGGCCCTGATG	927
Q	922	gaagcgctacccaaggtccgcgcgacatcgagaagtcgtcaagaagtgtcgatacttcgttaca	981
D	928	GAGGGATACCAATCCTCATCTTGGAGAGTGTGTCTCTGAGGCTGACATCTTTGTGATCC	987
Q	982	tgcacagaagaactgcgatatcatctctgttgacatagtgcagatgaagataaagct	1041
D	988	ACCACCGGAAACAGACACATCATCAATGATGACACATGACGAGAAATGAAGAACACGCC	1047
Q	1042	attgctcgtaacatcgccacttcgaatacgaagaattgatacagaatgcgcctcatgaat	1101
D	1048	ATTGCTCGAACAATTGTTGTAATCTTTGACATGAGATGACATGAAAGGCTTTGAGACCTAC	1107
Q	1102	ccaagcatcaagcacatcccaatccaagccagaatacgaacatgtggygaattcccaagatgc	1161
D	1108	CCTGTGTCTCAACCGCATCATCAATCAAGCCCAACATGACCGTGTGGCTTCCCGGAGACC	1167
Q	1162	caagct---atctcctcttctgtctgtagggccgcctcttaactgtgcttgcgtacaag	1218
D	1168	AAGACTGGCATATTGTTCTTCGTGAGGGTCGTCTAATGAACCTTGGATATGTCACATGGC	1227
Q	1219	caaccatcttcgtatagtcaatgttcattcaacaacaacagaacactcgcctgcacgaactc	1278
D	1228	CACCCAGTTTGTGATGTGCTATGTCGTCATTCACCAACACAGGTTATTGTGTGACTGATTTG	1287
Q	1279	taagaaag-----agagaaaatcgcgaagaaggttacaacacttcggaagcatctc	1332
D	1288	TGGAACTGAGAAAGCCAGTGGCAAGTATGTAGAAAGAAAGGTGTACCTTCCCCCAAGACACTT	1347
Q	1333	gatgaagaagtgcctgcgcctccaccccgagatcctcgaatgatacacttacaagaacttaca	1392
D	1348	GACGAGAAAGTGTGGCGCCCTCCACTTTGGGCAAGCTGGGCGCCAGGCTGACCAAGCTAC	1407
Q	1393	cagaagcaagctgactacatacgaagtlcagatgaaagtgltcctacaagtlcgtactac	1452
D	1408	AAGTCCCAAGTGTGACTACTATGTAGCAATCCCAATTGAGGGGTCTTACAAAGATGCGGCTTAC	1467
Q	1453	cgtaatta 1460	
D	1468	CGGTACTA 1475	

Search completed: August 17, 2002, 05:21:18

Job time: 12776 sec

---